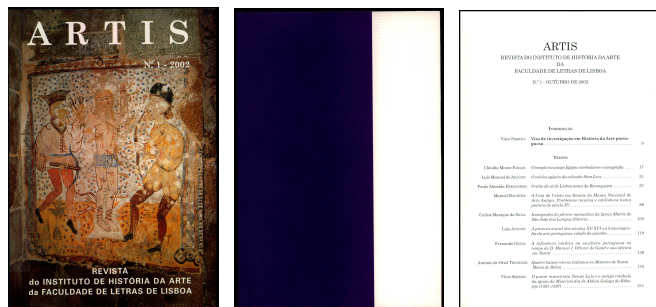


*O projecto de disponibilização de
Publicações Periódicas*



OUTUBRO 2009

UNIVERSIDADE DE LISBOA
FACULDADE DE LETRAS



BIBLIOTECA DIGITAL DA FLUL – 2ª FASE

*O projecto de disponibilização de
Publicações Periódicas*



Sónia Cristina Dias Teixeira

Mestrado em Ciências da Documentação e Informação (Biblioteconomia)

Trabalho Projecto orientado pelo **Professor Doutor Paulo Farmhouse
Alberto** e co-orientado pelo **Mestre Paulo Jorge Oliveira Leitão**

OUTUBRO 2009

RESUMO

O desenvolvimento das novas tecnologias de informação e comunicação conduziu a transformações significativas na forma como produzimos, disponibilizamos e utilizamos a informação. Face a estas mudanças, as bibliotecas tiveram que se adaptar não só às exigências de um público cada vez mais familiarizado com as novas tecnologias, mas também aos novos recursos que surgiram com o desenvolvimento tecnológico. Um bom exemplo desta nova realidade é o aparecimento das bibliotecas digitais.

É neste contexto que desenvolvemos um trabalho-projecto que versa sobre a temática das “bibliotecas digitais” e o seu impacto como novo meio de difusão e disponibilização da informação nas sociedades actuais.

Trata-se de um estudo que apresenta um projecto que permite aceder a uma colecção de objectos digitais através de uma biblioteca disponível em linha, efectuado e totalmente desenvolvido nas instalações da Biblioteca da Faculdade de Letras da Universidade de Lisboa (FLUL), no âmbito do projecto de Biblioteca Digital já existente.

O presente trabalho, “*Biblioteca digital da FLUL – 2ª fase. O Projecto de disponibilização de publicações periódicas*”, pretende apresentar um conjunto de soluções e princípios que permitam disponibilizar colecções digitais de uma forma eficaz e eficiente. Deste modo, procuramos responder, entre outras, às seguintes questões: como se constrói um projecto desta natureza; quais as boas práticas a seguir para a sua efectiva implementação; como desenvolver um projecto deste tipo *In-house*; que recursos técnicos utilizar. É neste sentido que trabalhamos com conceitos e normas que nos permitem compreender o que seria uma colecção digital de qualidade, quais os princípios fundamentais para uma digitalização eficiente, como implementar os decisivos metadados, como proceder no que respeita ao controlo de qualidade do material processado e, finalmente, como disponibilizar as colecções *On-line*. Este trabalho expressa ainda a importância e eficácia do uso de *software* livre neste contexto.

O resultado apresentado é o desenvolvimento de um projecto de disponibilização de publicações periódicas editadas pelos departamentos e centros de investigação da FLUL, apresentados agora via biblioteca digital.

PALAVRAS-CHAVE: biblioteca digital; colecção digital; digitalização; metadados; controlo de qualidade; disponibilização *On-line*; *software* livre.

ABSTRACT

The development of the new information and communication technologies has led to significant changes in how we produce, make available and use the information. Given these changes, libraries had to adapt not only to the needs of a new target audience more and more involved with the new technologies, but also to the new resources that emerged with the technological development. A good example of this new reality is the arising of the digital libraries.

In this context, we developed a project-work focused on the topic of “digital libraries” and its impact as a new way of information dissemination and availability in contemporary societies.

This is a study that presents a project that provides access to a digital collection through a digital library, made and fully developed in the facilities of the Library of the Faculty of Letters, University of Lisbon (FLUL) under the Digital Library project that already exist.

The present work, *“FLUL Digital Library – 2nd phase. The availability of periodical publications project*, aims to present a set of solutions and principles that enable a more efficient dissemination of the digital collections. Therefore, we intend to answer, among others, the following questions: how do we build such a project; what are the good practices to follow for its effective implementation; how to develop a project with these characteristics in-house; what technical resources should we use. Thus, we worked in this project with concepts and standards that allow us to understand what would be a high-quality digital collection, what are the fundamental principles for an efficient scanning, how to implement the critical metadata, how to check the quality of the processed material, and finally, how to make available these digital collections online. This work also expresses the importance and effectiveness of free software in this area.

The outcome of this work is the development of a project that provides access to periodicals published by the departments and research centres in FLUL presented now through its digital library.

KEYWORDS: digital library, digital collection, digitization, metadata, quality control, available online, free software

AGRADECIMENTOS

A elaboração deste trabalho não teria sido possível sem a colaboração, interesse e o apoio de várias pessoas a quem aqui expresso o meu agradecimento.

Em primeiro lugar quero agradecer o acompanhamento, apoio, empenho e compreensão dos meus orientadores, Doutor Paulo Farmhouse Alberto e Dr. Paulo Leitão. A disponibilidade e entrega por eles demonstrada foram fundamentais para o aperfeiçoamento deste trabalho.

Ao Dr. Pedro Estácio agradeço a oportunidade da realização de um estágio pós-laboral no seio da Biblioteca da FLUL e neste projecto em particular.

À Dr. Cristina Faria agradeço a disponibilidade e as acções de formação administradas, fundamentais para a compreensão e utilização dos diferentes programas.

À Ana Paula Canteiro. Este trabalho não teria sido realizado sem o apoio incondicional desta colega e amiga cuja ajuda e incentivo constantes foram preponderantes para o desenvolvimento e conclusão deste projecto.

Por fim, um agradecimento especial ao Nuno Pereira, pelo carinho e compreensão imperecíveis.

Obrigada!

ÍNDICE

CAP. 1 – INTRODUÇÃO	7
1.1 – Tema	7
1.2 - Objecto de estudo e objectivo do trabalho	7
1.3 – Motivação	8
1.4 - Metodologia do projecto	11
 CAP. II – AS BIBLIOTECAS DIGITAIS: CONTEXTUALIZAÇÃO E REVISÃO DA LITERATURA	14
2.1 – O conceito de biblioteca digital	14
2.2 - As bibliotecas digitais em universo universitário	16
2.3 – Constituição de uma biblioteca digital	18
2.3.1 – As colecções digitais	18
2.3.1.1 – Objectos digitais	20
2.3.2 – Recursos necessários	21
2.3.3 – Etapas para a efectiva constituição de uma biblioteca digital	23
 CAP. III – O PROJECTO DE DISPONIBILIZAÇÃO DE PUBLICAÇÕES PERIÓDICAS VIA BIBLIOTECA DIGITAL DA FLUL	32
3.1 – Caracterização da Instituição que acolhe o Projecto – FLUL	32
3.1.1 – A biblioteca da Faculdade de Letras	32
3.1.1.1 - A biblioteca digital	33
3.2 – O PROJECTO	34
3.2.1 - Público-alvo	35
3.2.2 - Selecção da colecção	35
3.2.3 - Identificação do problema	37
3.2.4 - Avaliação ou diagnóstico da situação – Análise SWOT	37
3.2.5 - Objectivos e Planeamento	45
3.2.6 - Implementação do projecto	45
3.2.6.1 – PROCESSO DE PRÉ-DIGITALIZAÇÃO	47
3.2.6.1.1 - Critérios de selecção dos exemplares	47
3.2.6.1.2 - O processo de levantamento	48
3.2.6.2 - O PROCESSO DE DIGITALIZAÇÃO	50
3.2.6.2.1 – Recursos técnicos	50
3.2.6.2.2 - Software utilizado – ScandAll	51

3.2.6.2.2 – Execução	51
3.2.6.3 – O PROCESSO DE IMPLEMENTAÇÃO DE METADADOS	53
3.2.6.3.1 - Software utilizado – PAPAIA	53
3.2.6.3.2 – Execução	54
3.2.6.4 - CONTROLO DE QUALIDADE DAS IMAGENS DIGITALIZADAS..	57
3.2.6.5 – O PROCESSO DE DISPONIBILIZAÇÃO	58
3.2.6.5.1 - O software utilizado – ContentE	58
3.2.6.5.2 – Execução	58
3.3. – RESULTADOS	61
CAP. IV – CONCLUSÕES	63
CAP. V – BIBLIOGRAFIA	66
ÍNDICE DE FIGURAS	
Figura 1 – Ficheiro “Listagem de Títulos” – <i>Revista Artis</i>	48
Figura 2 – Exemplo digitalização de capas e contra-capas	52
Figura 3 – Exemplo digitalização do conteúdo da obra	53
Figura 4 – Exemplo da normalização e estruturação das imagens – PAPAIA	56
Figura 5 – Tipo de índice utilizado na biblioteca digital da FLUL através do ContentE	59
Figura 6 – <i>Layout</i> da BND	59
Figura 7 – <i>Layout</i> da Biblioteca Digital da FLUL	60
ÍNDICE DE TABELAS	
Tabela 1 – Matriz de análise SWOT	38
Tabela 2 – Etiquetas TIFF utilizadas	54
ANEXOS	
ANEXO 1 – Cronograma da constituição do projecto	72
ANEXO 2 – Equipamentos técnicos	73
ANEXO 3 – O processo de levantamento	75
ANEXO 4 – O processo de digitalização	79
ANEXO 5 – O processo de implementação de metadados e nomeação de imagens	85
ANEXO 6 – O controlo de qualidade	95
ANEXO 7 – O processo de disponibilização – ContentE	98

CAP. I – INTRODUÇÃO

O presente trabalho-projecto denominado de **Biblioteca digital da FLUL – 2ª fase. O Projecto de disponibilização de publicações periódicas**, visa a obtenção do título de mestre em Ciências da Documentação e Informação.

Este projecto pretende tornar acessível em linha as revistas publicadas pelos Departamentos e Centros de Investigação da Faculdade de Letras da Universidade de Lisboa (FLUL), e tem vindo a ser desenvolvido na Biblioteca da FLUL, mais precisamente na Biblioteca Digital desta faculdade, desde Julho de 2008, por mim, enquanto estagiária e pelos restantes voluntários e funcionários da biblioteca.

1.1 – TEMA

A Internet veio alterar o modo como produzimos e disponibilizamos informação. Através das novas tecnologias da informação é possível publicar com mais facilidade e atingir uma diversidade de público maior. As bibliotecas, percebendo esta mudança, desenvolveram mecanismos de presença na Internet de que são exemplo os crescentes projectos de bibliotecas digitais que conduzem facilmente documentos do mundo analógico ao digital.

A temática central deste trabalho-projecto versa sobre a implementação de um projecto de biblioteca digital que visa a disponibilização *On-line* de uma colecção de publicações periódicas previamente seleccionadas para o efeito.

1.2 – OBJECTO DE ESTUDO E OBJECTIVO DO TRABALHO

O objecto de estudo desta dissertação é a biblioteca digital da FLUL e, mais particularmente, o novo projecto de disponibilização de periódicos nela implementado. Este projecto surge no seguimento do seu antecessor, *Património bibliográfico e documental da Faculdade de Letras da Universidade de Lisboa*, já existente na biblioteca digital, mas dedicado a espécies bibliográficas raras que existiam na instituição. Neste sentido, a implementação deste novo projecto correspondente ao que podemos chamar de uma 2ª fase da biblioteca digital da FLUL.

Nesta fase, optou-se por tornar acessível *On-line* as revistas publicadas pelos Departamentos e Centros de Investigação da FLUL que consistem nos periódicos: *Anglo-Saxónica*, *Ariane*, *Artis*, *Cadernos de Estudos Sefarditas*, *Cadernos de História de Arte*,

Cadmo, Clássica, Clio, Clio Arqueologia, Disputatio, Euphrosyne, Finisterra, Journal of Portuguese Linguistics, Op_Cit, Ophiussa, Phainomenon, Philosophica, Polifonia, Portugaliae-Histórica, Revista da Faculdade de Letras, Revista Lusitânia, Românica, Runa, Sinais de Cenas e Textos e Pretextos. A escolha deste universo de colecções surgiu devido a uma estratégia pensada pela instituição que, por um lado, pretendia dar mais visibilidade a outros objectos, diversificando a oferta e disponibilizando obras que primam pela sua actualidade e não somente pela sua raridade, e por outro lado, atingir novos públicos e dar continuidade ao projecto de biblioteca digital reduzindo os custos de investimento.

O objectivo deste trabalho é responder à pergunta: *Como implementar um projecto de disponibilização On-line de periódicos na já existente biblioteca digital da FLUL?* Pretende-se assim apresentar todo o processo que levará à disponibilização desta colecção na biblioteca digital.

1.3 – MOTIVAÇÃO PARA A ESCOLHA DO TEMA

No âmbito das bibliotecas digitais, especialmente das bibliotecas digitais universitárias, considerou-se pertinente apresentar um caso prático de um projecto de biblioteca digital.

A escolha deste tema justifica-se pela actualidade da temática aqui abordada. Assistimos a um constante desenvolvimento de projectos de bibliotecas digitais semelhantes, com maior incidência na área das bibliotecas universitárias, e tratando-se de uma área em crescimento e desenvolvimento, considerou-se importante abordá-la através de um estudo de caso real: a disponibilização *On-line* de revistas na Biblioteca digital da FLUL.

Ao longo da história as bibliotecas adoptaram diferentes modelos e adaptaram-se à sociedade, à comunidade onde estavam inseridas e da qual faziam parte, dando prioridade às motivações, interesses e necessidades dos seus públicos (público-alvo). Esta evolução é patente na oferta de serviços, que foram mudando em função das mutações do papel da biblioteca e do utilizador.

Presentemente, vivemos num novo paradigma que coloca o utilizador no centro de todos o sistema de informação, o utilizador actual não quer apenas “*ser servido*”, quer ter acesso a uma informação de qualidade e pertinente fornecida num período de tempo razoável e, se possível ainda, participar e produzir informação. A mesma transformação teve repercussões no papel do bibliotecário, neste paradigma, prestar um bom serviço passou a ser o mais importante, uma boa prestação de serviço seria a plena satisfação das

necessidades do utilizador o que implicava mudanças ao nível da própria disponibilização da informação.

O constante aparecimento de novas tecnologias da informação e o seu crescente uso nas nossas sociedades torna impreterível um olhar atento sobre elas, é imperioso perceber as novas potencialidades de algumas ferramentas e compreender o seu papel no que respeita à difusão da informação. As bibliotecas como “transmissoras de conhecimento” devem acompanhar estas inovações para que possam disseminar informação através das novas tecnologias informativas.

De facto, os tradicionais serviços das bibliotecas eram considerados insuficientes para acompanhar o utilizador actual, daí que tenham surgido serviços mais centrados no “novo” utilizador e que respondiam às suas necessidades actuais. É neste contexto que surgem os serviços de biblioteca digital, serviços que aliados às variadíssimas ferramentas disponibilizadas pela *Web 2.0* (wikis, blogues, etc) permitem à biblioteca transpor os seus “muros” e chegar aos utilizadores reais e potenciais. Os serviços de uma biblioteca digital vêm assim responder em parte a estas expectativas e ao novo modelo de disponibilização de informação.

O estudo “*College Students, Perceptions of Libraries and Information Resources*”, desenvolvido pelo OCLC, Online Computer Library Center¹, em 2005, já concluía que existia uma dissonância entre o ambiente e conteúdo que as bibliotecas providenciavam e o ambiente e conteúdo que os utilizadores /consumidores queriam e usavam. Segundo a OCLC (2006), tem sido cada vez mais difícil caracterizar e descrever o propósito das bibliotecas porque a relação entre a informação, o utilizador e o conteúdo mudou e continua a mudar.

Este estudo, efectuado com o intuito de compreender a percepção dos estudantes universitários face às bibliotecas e fontes de informação, foi realizado em seis países distintos (Austrália, Canada, Índia, Singapura, U.K e E.U.A) e permitiu obter uma visão relativamente global da posição dos estudantes face a esta problemática e entender o novo paradigma centrado no utilizador, que as bibliotecas já enfrentam.

Os resultados indicam que 89% das pesquisas de informação efectuadas pelos estudantes começam por um motor de busca e apenas 2% dos estudantes escolhe o sítio da biblioteca para esse efeito, mais, apenas 10% dos estudantes afirma que o sítio da biblioteca preenche as suas necessidades de informação o que é revelador da dificuldade

¹ A OCLC é uma organização não governamental que promove a cooperação entre bibliotecas de todo o mundo. Foi criada em Ohio em 1967 e actualmente usam os seus serviços mais de 54.000 bibliotecas e 109 países.

das bibliotecas acompanharem os novos públicos. Certamente que quatro anos volvidos sobre este estudo, se registaram melhorias nos sítios *Web* das bibliotecas. Mas será o suficiente? Segundo o estudo, “*os estudantes estão a ler menos e a usar a biblioteca menos desde que começaram a usar a Internet*” (OCLC, 2006, p.3-19)

O estudo revela também que a maioria dos estudantes vê a biblioteca apenas como uma instituição “fornecedora de livros”. Segundo o estudo, uma razão potencial para a desconexão entre a percepção dos utilizadores (bibliotecas = livros) e a percepção dos bibliotecários (biblioteca = um variado conjunto de produtos e serviços) é uma falha da educação do utilizador. De facto, existe um grande desconhecimento no que respeita aos variados serviços que as bibliotecas disponibilizam, contudo, será correcto afirmar que isso é uma falha dos estudantes ou será antes uma falha das bibliotecas na divulgação dos seus serviços?

No que respeita à qualidade da informação e confiança, poderíamos pensar que claramente os estudantes apontariam as bibliotecas como “centros de saber” onde a informação é mais credível do que num simples motor de busca na Internet, no entanto, o estudo não revela esta ideia. Os estudantes praticamente não fazem distinção entre a qualidade da informação e a confiança que sentem ao pesquisarem numa ou noutra. Nitidamente as bibliotecas têm vindo a perder público e estão a ser facilmente ultrapassadas pelos novos meios de divulgação da informação através da Internet cujo exemplo paradigmático é o motor de busca.

Tendo em conta estes dados, como ultrapassar esta tendência tão desfavorável para as bibliotecas? É importante que as bibliotecas apresentem um serviço que traduza a realidade dos nossos dias em que a vertente tecnológica é quase inerente ao indivíduo e é o principal mecanismo de procura de informação do mesmo. As bibliotecas têm que estar presentes *On-line*, indo ao encontro das necessidades e preferências dos utilizadores. Contudo, não basta dispor de um sítio *Web*, este deve ser quase diariamente actualizado, tem de permitir que o utilizador participe, deixe a sua opinião, e ainda oferecer uma série de serviços apelativos e que respondam às necessidades dos utilizadores.

É neste contexto que surge a biblioteca digital, uma parte da biblioteca que se encontra *On-line*. Actualmente, não é suficiente disponibilizar apenas um catálogo *On-line* por exemplo. É necessário ir mais além, disponibilizar informação no instante, aproximando-se ao máximo do nível de satisfação que o utilizador obtém quando pesquisa informação na Internet, ou seja, informação no momento, simples de pesquisar e de resultado célere. Obviamente que uma biblioteca digital não pode ser uma versão *On-line* da “biblioteca mãe”, mas pode ser um canal bastante útil para apresentar as suas melhores colecções e estar disponível mundialmente.

É por este motivo que retratamos um caso de biblioteca digital neste trabalho, por considerarmos ser um mecanismo importante que responde em parte às actuais necessidades de informação dos utilizadores. Assim sendo, podemos dizer que a oferta de bibliotecas digitais nas universidades como serviço de informação é uma questão de sobrevivência, *“Para serem competitivos e se adaptarem às novas necessidades dos utilizadores, os serviços de informação das universidades têm de oferecer o que procuram os novos sistemas de docência, aprendizagem e hábitos dos utilizadores.”* (FERRER SAPENA, [et al.], 2005, p. 13). Têm de se tornar nas chamadas bibliotecas híbridas, isto é, bibliotecas que oferecem a utilização tradicional de suportes em papel aliada a uma vasta gama de suportes digitais.

A motivação para problematizar esta temática surge, portanto, no sentido de ir ao encontro deste novo paradigma de disponibilização de informação que actualmente presenciamos, mostrando como se pode construir e disponibilizar um projecto de biblioteca digital.

1.4 – METODOLOGIA DO PROJECTO

Naturalmente, a abordagem desta temática das bibliotecas digitais em geral, e dos projectos de bibliotecas digitais universitárias em particular, foi precedida de uma revisão da literatura relevante sobre métodos de desenvolvimento de bibliotecas digitais, sobre projectos de bibliotecas digitais universitárias, a implementação de metadados, as boas práticas da digitalização e disponibilização. Deste modo, e partindo desta análise da literatura, serão definidos uma série de critérios e analisados outros já publicados de modo a serem aplicados ao projecto. Estes critérios apresentados servirão de modelo para perceber quais serão as práticas e normas mais indicadas a implementar neste caso.

A metodologia de investigação utilizada neste trabalho foi uma *metodologia projecto*, seguindo parte das directrizes da obra *“Guia metodológica para la implantación de una biblioteca digital universitaria”* de Antonia Ferrer Sapena [et al.] (2005), que indica procedimentos a aplicar na realização de um projecto de biblioteca digital. No entanto, dado que este guia está direccionado para a construção de raiz de bibliotecas digitais universitárias, resolvemos implementar algumas alterações a esta metodologia de forma a orientá-la para as necessidades particulares do nosso projecto que se integra numa biblioteca digital já existente.

Sendo que um projecto é um conjunto de actividades executadas por uma equipa no sentido de atingir um determinado objectivo, utilizando diversos recursos e uma estratégia adequada, a *“Metodologia do Trabalho de Projecto consiste na adopção de um conjunto*

de procedimentos, técnicas e instrumentos com vista a atingir os objectivos do projecto” (RAMOS, 2008, p. 7). No fundo, a metodologia do projecto constitui a ponte entre teoria e prática porque possui como suporte o conhecimento teórico que será posteriormente utilizado na implementação de um trabalho concreto e real.

O presente trabalho é composto por cinco capítulos. Os capítulos base, introdução, conclusão e bibliografia e dois capítulos de desenvolvimento, designadamente o capítulo II e capítulo III. O segundo capítulo possui um cariz introdutório, faz um enquadramento sobre a temática das bibliotecas digitais, definindo o conceito, abordando a pertinência desta temática no mundo actual e apontando para a questão das bibliotecas digitais no contexto universitário. Ainda neste capítulo referimos os princípios, as normas e procedimentos para a constituição de uma biblioteca digital de qualidade.

O terceiro capítulo começará por apresentar a instituição que acolhe o projecto, a biblioteca da Faculdade de Letras da Universidade de Lisboa, seguido da abordagem de todas as vertentes relacionadas com a implementação do projecto em si.

Para a execução deste projecto foram definidas 6 etapas, baseadas na metodologia já referida, que permitiram obter uma estrutura lógica de desenvolvimento do projecto.

1. **Definição público-alvo:** definição da população para a qual se destina o projecto.
2. **Seleção da colecção:** justificação da escolha desta colecção para fins de disponibilização *On-line* e caracterização da mesma.
3. **Identificação do problema:** identificação da problemática que o projecto pretende resolver.
4. **Avaliação ou diagnóstico:** realização de um diagnóstico da situação através de uma análise SWOT (*Strengths, Weaknesses, Opportunities e Threats*)
5. **Objectivos e planeamento:** definição dos objectivos do projecto e o planeamento efectuado para a realização das actividades.
6. **Implementação do projecto:** apresentação das actividades desenvolvidas para a execução do projecto – levantamento dos exemplares e os processos de digitalização, de implementação de metadados, de controlo de qualidade e de disponibilização dos conteúdos.

A lógica desta ordenação de etapas é facilmente explicável. A biblioteca tendo consciência do público-alvo real e potencial de um projecto desta natureza seleccionou a colecção que pretendia disponibilizar via biblioteca digital. Entretanto, deparou-se com uma questão – como implementar este projecto em concreto enquadrando-o no projecto já existente de biblioteca digital? Para solucionarmos este problema fizemos um diagnóstico

da situação através do método “análise SWOT” que nos permitiu identificar os principais aspectos que caracterizam a biblioteca digital, mais precisamente este projecto, para posteriormente definirmos objectivos e planear as actividades a desenvolver. Primeiramente fizemos o levantamento dos exemplares da colecção a digitalizar e por fim demos início à parte prática de execução do projecto através dos vários processos nele envolvidos.

No que respeita à execução do processo de digitalização, serão apresentadas as opções tomadas, decorrentes das boas práticas implementadas na instituição, resultantes da experiência empírica do projecto anterior. O mesmo se passa no que concerne ao controlo de qualidade, serão apresentados os critérios definidos para que este controlo seja o mais rigoroso possível, seguindo as políticas da instituição para o efeito e as normas internacionais definidas. De igual forma as fases de produção de metadados e da disponibilização serão desenvolvidas baseando-se nas boas práticas da instituição e da Biblioteca Nacional Digital (BND).

A primeira revista a ser tratada para fins de disponibilização em acesso livre e que serviu de base à implementação deste projecto, foi a revista *Artis: Revista do Instituto de História da Arte da Faculdade de Letras de Lisboa*, logo, serão os seus exemplos que surgiram no trabalho aquando da apresentação do projecto. Referir-nos-emos a ela apenas por *Artis*.

CAP. II – AS BIBLIOTECAS DIGITAIS: CONTEXTUALIZAÇÃO E REVISÃO DA LITERATURA

O surgimento das bibliotecas digitais resulta da evolução da própria tecnologia, desde a criação do primeiro computador, na década de 1940, até à invenção da *World Wild Web* por Tim Berners-Lee em meados de 1990.

As primeiras tentativas de armazenar informação digital em computadores, no final da década de 60 foram frustradas devido às limitações técnicas e também ao elevado custo dos computadores e dos *interfaces* dos utilizadores. As décadas de 70 e 80 corresponderam ao aparecimento das redes de comunicação de dados que culminam com a consulta remota da base de dados *On-line*, nomeadamente os OPAC (Online Public Access Catalog); contudo, quanto à produção e armazenamento de informação digital não se registou uma grande evolução porque o processo continuava a ser dispendioso.

Durante a década de 90, esta situação inverteu-se e estas barreiras começaram a ser ultrapassadas devido a uma série de desenvolvimentos tecnológicos que permitiram a aquisição de computadores mais económicos estimulando a prestação de serviços *On-line*. Este desenvolvimento tecnológico despoletou mudanças significativas, o armazenamento electrónico tornou-se mais barato do que o papel e neste sentido, o custo de armazenar documentos em CD-ROM era inferior a armazenar livros numa biblioteca.

Actualmente, testemunhamos o resultado destas transformações. É muito mais económico armazenar informação num computador do que numa biblioteca, ainda que, armazenar grandes objectos como vídeos, extensas colecções de imagens, ou sons de alta qualidade seja ainda dispendioso. Contudo, com a constante evolução tecnológica prevê-se que no espaço temporal de apenas uma década o equipamento seja tão acessível que a questão do preço não será um factor a considerar na tomada de decisão na criação ou não de uma biblioteca digital.

2.1 – O CONCEITO DE BIBLIOTECA DIGITAL

O conceito base aqui tratado é a Biblioteca Digital, mas como o podemos definir? Em que consiste? Uma biblioteca digital é um sistema de informação constituído por uma colecção de objectos digitais, originalmente digitais ou resultantes da digitalização de objectos analógicos, organizados. As suas vantagens primárias para o utilizador são o acesso remoto através de um computador com ligação em rede, a utilização simultânea dos

seus conteúdos por diversos utilizadores e ainda a possibilidade de utilizar de forma integrada diferentes registos de informação como o texto, o som e imagem.

É correcto dizer que estas bibliotecas são colecções organizadas de informação digital, mas segundo a *Digital Library Federation* uma biblioteca digital é muito mais que uma colecção de textos e outros objectos digitalizados, uma biblioteca digital é uma organização que “*disponibiliza recursos, incluindo pessoal especializado, para seleccionar, estruturar, oferecer acesso intelectual para interpretar, distribuir, preservar a integridade e assegurar a persistência ao longo do tempo de colecções digitalizadas de modo a que estejam prontamente e economicamente disponíveis para o uso de uma determinada comunidade ou conjunto de comunidades.*” (Digital Library Federation, 2001)

As bibliotecas digitais vêm responder à ideia de biblioteca “sem muros”, uma vez que eliminam as barreiras físicas e a distância, factores que desde sempre limitaram o sector das bibliotecas físicas. Todavia, estas bibliotecas sofrem outros tipos de limitações, nomeadamente a nível da sua temática e do conteúdo. Como delimitar os conteúdos a disponibilizar sendo que não é possível digitalizar tudo? Esta é uma questão que nos remete para a definição dos critérios de selecção da colecção que devem ser bem fundamentados pelas respectivas bibliotecas.

Segundo Arms, uma “*biblioteca digital é uma colecção organizada de informação, com serviços associados, onde a informação é armazenada em formatos digitais e acessível através de uma rede*”. (2001, p. 2) Ressalva ainda que o que se destaca nesta definição é o termo “colecção organizada de informação”, isto porque, um conjunto de dados disponível *On-line* não é por si só uma biblioteca digital, somente o será se esses mesmos dados se encontrarem organizados.

É facto notório, não obstante, que a maior parte das bibliotecas digitais que conhecemos não existem na íntegra. O que existe são as chamadas bibliotecas híbridas, termo usado para descrever bibliotecas que contêm uma mistura de recursos tradicionais de uma biblioteca (ex: livros, revistas) e de recursos electrónicos (ex: audiobooks, jornais electrónicos e outros conteúdos digitais).

Podemos dizer que, no que respeita às bibliotecas digitais, existem cinco conceitos basilares: colecção, tecnologia, pessoas, acesso e preservação. De facto, uma biblioteca digital deve ser constituída por uma *colecção* ou conjunto de colecções que têm uma estrutura e um sentido; tem de possuir um sistema técnico, ou seja, uma *infra-estrutura tecnológica* sem a qual não subsiste; tem de ter *pessoal* especializado que a gere e tem de *preservar* a integridade dos objectos disponibilizando o *acesso* aos mesmos.

Mas porquê digitalizar as colecções? Quais os benefícios e as desvantagens desse processo? Se é verdade que as bibliotecas digitais apresentam um conjunto de vantagens face às chamadas bibliotecas tradicionais, também apresentam algumas desvantagens ou limitações, embora se pense que estas limitações, quando comparadas, sejam muito inferiores.

VANTAGENS

- a) Permitem o acesso à distância;
- b) Funcionam 24 horas por dia;
- c) Permitem o acesso simultâneo de vários utilizadores;
- d) Podem conjugar diferentes medias de informação (texto, imagem, som);
- e) Possibilitam a preservação dos documentos;
- f) Facilitam o acesso a pessoas com deficiência;
- g) Permitem o acesso em linha a outras fontes de informação externas;
- h) Os custos inerentes são cada vez mais reduzidos;
- i) Podem facilitar processos de aprendizagem e educação.

DESVANTAGENS

- a) Eventuais problemas relacionados com o direito de propriedade (Direitos de Autor);
- b) O acesso, o armazenamento e a preservação das obras digitalizadas estão dependentes do desempenho das tecnologias;
- c) Excesso de informação que pode criar redundância;
- d) O acesso depende dos preços das ligações à Internet.

A informação apresentada em formato digital ultrapassa as barreiras espaciais ou temporais e o original pode ser copiado infinitamente sem que disso resulte a sua degradação. Contudo, talvez uma das características/factores mais importantes da informação digital seja o facto de esta poder estar ligada a uma infinidade de conteúdos *Web*, tanto a nível local como global e de permitir a pesquisa e a navegação de uma forma simples e de resultados instantâneos.

2.2 - AS BIBLIOTECAS DIGITAIS PORTUGUESAS EM UNIVERSO UNIVERSITÁRIO

A tecnologia desempenha actualmente um papel fundamental nas organizações, ela permite o desenvolvimento de organizações digitais, que ultrapassam obstáculos geográficos e temporais na prestação de serviços e permitem novos níveis de colaboração.

Estes serviços prestados através de bibliotecas digitais, por exemplo, começam a ganhar cada vez mais peso nas bibliotecas e começam a ser em Portugal espaços que possibilitam ao utilizador uma série de serviços válidos e úteis.

Na verdade, todas as bibliotecas digitais portuguesas advêm de “instituições híbridas” ou seja, um misto de biblioteca tradicional e digital cuja missão é gerir recursos culturais tanto analógicos como digitais, e é esta a tendência das bibliotecas universitárias.

As bibliotecas universitárias têm tido um papel preponderante na propagação de projectos de biblioteca digital em Portugal. Destacamos neste trabalho os exemplos da Universidade de Coimbra e da Universidade da Beira Interior que apresentam um trabalho que resulta em diversas bibliotecas digitais com diferentes conteúdos e áreas temáticas. A Universidade de Coimbra regista neste momento quatro projectos (a Biblioteca Geral Digital; a Biblioteca Digital da Faculdade de Direito; a Biblioteca Digital de Botânica e o Centro de Documentação 25 de Abril). A Universidade da Beira Interior por sua vez dispõe de cinco (a BOAL – Biblioteca *On-line* Áudio de Literatura; a BOCC – Biblioteca *On-line* de Ciências da Comunicação; a LUSOSOFIA – Biblioteca *On-line* de Filosofia; a THESIS – Repositório de Publicações e a UBILETRAS – Revista *On-line* do Departamento de Letras).

Destacamos também evidentemente a Universidade de Lisboa com a Biblioteca Digital da Faculdade de Letras e a Universidade de Aveiro com a BAES – Biblioteca Aberta do Ensino Superior, uma biblioteca digital com conteúdos acessíveis à comunidade de utilizadores com necessidades especiais, que resulta de um projecto em parceria entre as Universidades do Porto, Minho, Aveiro, Coimbra, Évora, Lisboa - Reitoria, Faculdades de Letras e de Ciências da Universidade de Lisboa e Trás-os-Montes e Alto Douro. Distinguimos ainda a importância destes projectos de parceria efectuados entre as universidades que proporcionam a partilha de recursos e conteúdos.

Na verdade, é actualmente sinónimo de modernidade uma biblioteca ter o seu projecto de biblioteca digital, atesta que a biblioteca está activa e empenhada em chegar aos seus públicos e ao mesmo tempo proporciona à instituição uma visibilidade notável. Um projecto desta natureza pode funcionar como uma estratégia de marketing para dar a conhecer mundialmente as melhores colecções daquela biblioteca em particular.

Os repositórios das universidades, por exemplo, desempenham um papel fundamental na partilha de conteúdos digitais entre as várias universidades e podem ser entendidos como bibliotecas digitais que possibilitam o acesso à produção científica universitária.

Nos repositórios podemos encontrar diversos tipos de documentos, em formato digital, resultantes da produção científica das instituições como teses de doutoramento,

teses de mestrado, artigos em publicações periódicas nacionais e internacionais, relatórios técnicos, artigos em actas de conferências nacionais e internacionais, capítulos de livros, entre outros.

Exemplo de um repositório onde se podem conjugar todas estas colecções é o RCAAP (Repositório Científico de Acesso Aberto de Portugal). É um repositório que possui duas funções. Por um lado, é utilizado por instituições de ensino e investigação produtoras de literatura científica que não têm um repositório próprio e, por outro lado, funciona também como um canal centralizador que disponibiliza o acesso aos repositórios universitários já existentes.

2.3 – A CONSTITUIÇÃO DE UMA BIBLIOTECA DIGITAL

Em termos gerais, uma biblioteca digital modelo será aquela que satisfaz o seu público quer ao nível do conteúdo, quer ao nível do acesso à informação. Para além disto, ao nível técnico, uma biblioteca digital terá de ser capaz de possibilitar e contribuir para a preservação dos seus documentos a longo-prazo.

Desta forma, do ponto de vista do conteúdo, uma biblioteca digital deve conter colecções que respondam às necessidades de um público cada vez mais diversificado e exigente, e a informação deverá estar disponível em texto formatado ou em imagem digitalizada. Do ponto de vista da lógica organizativa, deve apresentar uma estrutura ordenada que seja facilmente descortinada pelo utilizador. Não será de somenos importância disponibilizar uma informação de suporte ao utilizador através da apresentação de um “welcome” com a informação sobre a colecção, ajudas, contactos e os direitos e condições de utilização da colecção, por exemplo. Sobre o prisma das formas de acesso, a biblioteca digital deve ter um *interface* público para disponibilizar o acesso à sua colecção e conter instrumentos que permitam a “navegação” através da informação, como o *browsing* (percorrer), o *search* (pesquisa) que devem permitir a pesquisa por assuntos, autores e título. Será também importante que o utilizador consiga chegar aos conteúdos por meio da pesquisa sobre a metainformação, através de um catálogo, e por OCR (*Optical Character Recognition*).

2.3.1 – AS COLECÇÕES DIGITAIS

As bibliotecas digitais são constituídas por colecções digitais, isto é, por conjuntos organizados de objectos ou imagens digitais com determinadas características comuns.

Como planear uma colecção digital que constituirá a biblioteca digital? Para criarmos uma colecção digital com fundamento e lógica temos de responder a uma série de perguntas que nos ajudam a definir o trajecto a seguir: qual é o objectivo da colecção, qual é o fim que quer atingir? Qual é o seu público-alvo? Para que fins vai a colecção ser utilizada? A resposta a estas questões ajuda-nos a definir o propósito da colecção que pretendemos disponibilizar e, consequentemente, os objectos que a constituirão.

No sentido de criar colecções digitais mais completas tanto ao nível do conteúdo como ao nível da qualidade do mesmo, muitas instituições criaram políticas internas de digitalização, selecção e disponibilização, no sentido de definir as chamadas “*best practice*” da instituição, isto é, as boas práticas a utilizar aquando de um projecto desta natureza para facilitar o processo de “como fazer” respeitando uma série de parâmetros. A *Colômbia University Library*, a *Oxford University*, *Harvard* e a *University of California* são exemplos disso. Instituições que criaram estes documentos internos com normas adequadas à sua realidade e que disponibilizam *On-line* para servir de referência a outras bibliotecas. A University of Maryland (E.U.A), por exemplo, criou um documento deste tipo denominado de *Best practice guidelines for digital collections at University of Maryland Libraries* em 2006 e cuja 2.^a edição surgiu em 2007. Este documento descreve o que a instituição entende como “boas práticas” a seguir para desenvolver colecções digitais de qualidade, é um documento muito completo e tem a mais valia de poder ser utilizado por outras bibliotecas.

Neste âmbito, destacamos também o trabalho desenvolvido pela NISO (National Information Standards Organization), uma associação sem fins lucrativos credenciada pelo American National Standards Institute (ANSI). Esta organização identifica, desenvolve e publica normas técnicas para a gestão da informação, um campo em constante mudança e cada vez mais digital. É uma referência mundial na normalização e não sendo uma ISO (International Organization for Standardization) é de facto uma referência importante.

Em 2007 publicou a 3.^a edição do documento *A Framework of Guidance for Building Good Digital Collections*, um guia que apresenta recomendações de “boas práticas” a adoptar na construção de uma colecção digital. Destacamo-lo neste trabalho precisamente por considerarmos que é uma ferramenta muito útil para a regulamentação e planeamento de uma boa colecção digital.

Segundo este documento, uma colecção digital “*consiste em objectos digitais que são seleccionados e organizados para facilitar a sua descoberta, acesso e uso. O que forma a colecção é a conjugação dos elementos objectos, metadados e a interface utilizada.*” (NISO, 2007, p. 4)

O guia da NISO determina critérios para uma boa organização dos trabalhos em relação a quatro tipos de entidades: colecções, objectos, metadados e iniciativas ou projectos para criar e gerir colecções. O uso destas práticas recomendadas pela NISO é opcional. Pode ser usado como indicado ou modificado por quem as utiliza no sentido de atender às necessidades específicas da sua colecção. Estas linhas orientadoras são, de facto, uma referência extremamente importante para qualquer projecto deste tipo. Utilizaremos estas recomendações como base para definirmos uma colecção digital de qualidade no projecto que apresentaremos.

Os objectos que constituem as colecções digitais têm de ter obrigatoriamente um determinado conjunto de atributos que lhe conferem qualidade. Segundo a NISO (2007), esses atributos são:

1. Um bom objecto existe num formato que suporta o seu uso actual e futuro pretendido;
2. Um bom objecto é preservável, isto quer dizer que o objecto será acessível apesar das mudanças da tecnologia. (estratégias: migração, emulação; formatos não-proprietários);
3. Um bom objecto tem um significado e é útil fora do seu contexto local;
4. Um bom objecto será nomeado com um identificador persistente, globalmente único que pode ser reportado (resolved) à actual morada (localização) do objecto;
5. Um bom objecto pode ser autenticado;
6. Um bom objecto tem associado metadados.

2.3.1.1 - OBJECTOS DIGITAIS

O objecto que constitui uma colecção digital é a imagem digital normalmente resultante de um processo de digitalização. Mas em que consiste essa imagem? Uma imagem digital é uma representação através de um conjunto finito de pontos definidos por valores numéricos, formando uma malha onde cada ponto é um pixel.

Um Pixel, da junção dos termos *Picture* e *Element*, é o ponto que dá forma a uma imagem digital, sendo que o conjunto de milhares de pixels forma uma imagem inteira. Este processo gera-se da seguinte forma: a cada pixel é atribuído um valor tonal de preto, branco, escala de cinzentos ou cor, representado por um código binário de zeros e uns. Os dígitos binários, bits, de cada pixel são armazenados por um computador numa dada sequência de forma a compor a imagem. Estes bits podem ainda ser comprimidos para tornar uma imagem menos pesada.

A resolução representa o nível de detalhe da representação da informação e é expressa em dpi (dots per inch) pontos por polegada, ou em ppi (pixels per inch) pixel por polegada, sendo que estes determinam o nível de representação de detalhe da imagem.

Já a profundidade de cor (bit depth) refere-se à cor ou à escala de cinzentos de um pixel e é determinada pelo número de *bits* utilizado para definir cada *pixel* nas imagens digitais. Um pixel com 8 bits por cor resulta numa imagem de 24 bits (8 bits multiplicados pelas cores RGB - *red*, *green* e *blue*, o que origina uma imagem de 24 bits).

Quando se captura electronicamente imagens por *scanners* ou câmaras digitais, estas são transferidas para um computador onde são armazenadas num determinado formato, sendo os mais comumente utilizados, os formatos TIFF, JPEG e GIF. Estes formatos descrevem as características que queremos que cada imagem possua. Cada formato apresenta diferentes vantagens e desvantagens e diferentes níveis de compressão, nesse sentido, a sua escolha depende do uso que queremos dar às imagens.

O tamanho de ficheiros, das imagens, depende da altura e largura da imagem, da sua qualidade e do formato no qual são gravados. Por vezes, um sistema de compressão é utilizado para reduzir o tamanho da imagem, contudo, esta técnica é apenas utilizada nas imagens que servirão de cópias para fins de disponibilização com o intuito de se tornarem mais leves. As imagens de arquivo, denominadas “*master*”, são digitalizadas em formato TIFF. Uma imagem *master* resulta da digitalização de alta qualidade de um objecto (imagem/texto), sendo que o objectivo é reproduzir a realidade analógica da maneira mais precisa possível, logo, são reproduzidas em formatos que não possuem compressão e com altos níveis de qualidade.

De facto, conclui-se que para constituir uma biblioteca digital de qualidade é necessário possuir uma boa colecção digital e isso passa desde logo pela constituição dos objectos digitais. O termo “boa” neste contexto “*exige interoperabilidade, re-uso, persistência, verificação, documentação, e sustentação dos direitos de propriedade intelectual.*” (NISO, 2007, p.1)

Contudo, a criação de uma colecção digital com qualidade deve passar por várias etapas e obedecer a requisitos técnicos e princípios de qualidade que serão de seguida explanados.

2.3.2 - RECURSOS NECESSÁRIOS

a) INFRA-ESTRUTURA E TECNOLOGIA

O desenvolvimento de uma biblioteca digital pressupõe a criação de infra-estruturas básicas que nos assegurem a sua eficiência. É necessário possuir uma infra-estrutura

tecnológica que se coadune com as necessidades de um projecto desta natureza, isto é, *hardware* e *software* que permitam a constituição e preservação da colecção e o acesso à mesma.

Num projecto de biblioteca digital são preponderantes as condições de *hardware* (servidores, discos rígidos, monitores, *scanners*,...) e *software* que suportam a colecção, desta forma, devem possuir uma grande capacidade de armazenamento e funcionalidade, respectivamente. Ambos *devem não só atender às necessidades presentes mas estar perfeitamente preparados para as necessidades futuras ou para se adaptarem a elas (...) o software deve ser adaptável*” (FERRER SAPENA, et.al., 2005, p. 83). O *Hardware* e o *Software* devem ser compatíveis para evitar erros, o mesmo se passa com o sistema operativo escolhido, sendo fundamental determinar se é compatível ou suporta os programas que queremos utilizar.

Os dispositivos de saída e entrada também desempenham um papel fundamental neste campo, eles são os responsáveis pelos dados que inserimos (*inputs*) e pelos *outputs* que obtemos através de uma operação informática.

Os monitores são dispositivos de saída indispensáveis no desenvolvimento de um projecto desta natureza, devem possuir uma boa resolução para que a visualização das imagens e das eventuais imperfeições seja correctamente efectuada. Os *scanners* como dispositivos de entrada são também preponderantes, deles depende em grande parte o sucesso do projecto, como tal, devem ser considerados os parâmetros: resolução óptica, profundidade de cor, rapidez, conectividade e as funcionalidades do *software* associado.

b) RECURSOS HUMANOS

Segundo a obra, *Guía metodológica para la implementación de una biblioteca digital universitaria* (FERRER SAPENA, et.al., 2005, p. 90), os recursos humanos que devem compor o grupo de trabalho num projecto de biblioteca digital são os seguintes:

- *Pessoal auxiliar encarregado da preparação do material para a digitalização;*
- *Pessoal auxiliar para a digitalização das imagens, textos, etc;*
- *Pessoal técnico para o tratamento de textos e sua colocação em rede;*
- *Pessoal informático;*
- *Coordenador do projecto;*
- *Pessoal encarregado da sua manutenção;*
- *Pessoal encarregado dos serviços criados;*
- *Pessoal especializado em análise da informação.*

Obviamente que o ideal seria possuir um técnico de cada uma destas áreas a trabalhar em conjunto, mas não podemos esquecer que nem todas as bibliotecas podem

dispor destes recursos. No entanto, existe sempre a possibilidade de seleccionar pessoal polivalente que muitas vezes combina em si várias destas competências.

c) RECURSOS FINANCEIROS

A melhor forma de desenvolver um projecto de digitalização é conhecer todos os aspectos que envolvem o processo de digitalização tanto a nível técnico, que equipamentos usar, que normas seguir, entre outros, como a nível administrativo, quantas pessoas contratar para cada tarefa, gerir horários de voluntariado, caso se aplique, gerir formações, etc. Assim sendo, um projecto de digitalização apresenta invariavelmente custos de várias ordens. Estes custos têm de ser previamente pensados e discutidos de modo a assegurar que o projecto não fique comprometido durante a sua execução por falta de cumprimento de um ou vários pontos preponderantes. Segundo Lorna Hughes (2004), há que ter em conta os seguintes custos:

- a) *Custos materiais* que incluem todo o equipamento necessário para materializar o projecto, computadores, scanners, software, formação, serviços de “guilhotinagem”, discos externos, entre outros;
- b) *Custos organizacionais* como os custos de postos de trabalho relacionado com o espaço de escritório, acesso à rede e espaço de rede disponível;
- c) *Custos de mão-de-obra* que normalmente são os mais dispendiosos;
- d) Outros custos associados ao acesso e preservação do projecto a longo prazo.

2.3.3 – ETAPAS PARA A EFECTIVA CONSTITUIÇÃO DE UMA BIBLIOTECA DIGITAL

a) A SELECÇÃO DO MATERIAL

Uma vez que digitalizar toda a colecção que constitui o espólio de uma biblioteca é inexequível, há que definir concretamente a parcela daquele universo que será útil digitalizar e disponibilizar *On-line*, por outras palavras, será desejável a definição de uma política que estabeleça os critérios de selecção.

Esta política é influenciada por vários factores tais como: o interesse público da obra, os direitos de autor, obras que são frequentemente requisitadas, obras que ainda não sofreram um processo de digitalização, entre outros. Segundo Lorna Hughes (2004) “*o principal critério para a digitalização deve ser: o conteúdo informacional dos materiais (qual é o valor intelectual da colecção?); a procura dos materiais (como são usados, e em que frequência, por que tipo de utilizadores?) e as condições do original (são frágeis ou estão danificados e é viável a digitalização?)*” (HUGHES, 2004, p. 36).

Todos estes factores conjugados levam à selecção de um pequeno grupo de obras, uma vez que, apenas um pequeno grupo conseguirá ter aval positivo em todas as vertentes que apontamos. Segundo a mesma autora, *“A digitalização pode dar uma nova vida a recursos importantes e permitir o acesso a uma comunidade externa de utilizadores. Como tal, a decisão de digitalizar deve ser baseada no conhecimento do que é único acerca da informação das colecções.”* (HUGHES, 2004, p. 32).

Segundo o documento *Selection Criteria for Preservation Digital Reformatting* (2002) da Biblioteca do Congresso, a selecção dos materiais para preservação digital é baseada no valor, uso, condição, características do original e adequação de reproduções digitais para uso e acesso. Assim sendo, é notório que antes de se começar um projecto de digitalização como o que aqui retratamos, é necessária uma avaliação cuidadosa do conteúdo intelectual da colecção.

b) A DIGITALIZAÇÃO

Findo o processo de selecção da colecção a digitalizar, passamos para a etapa seguinte da digitalização propriamente dita. Esta é talvez a fase mais importante de todo o processo de implementação de uma colecção digital e consequente biblioteca digital.

A digitalização é o processo de criação de ficheiros digitais pela conversão de material analógico. Neste processo, o conteúdo analógico é convertido numa sequência de 0 e 1 e colocado num código binário para ser lido por computador. A cópia digital resultante pode ser depois classificada como material digital e o principal desafio envolvido nesta questão é preservar o acesso ao mesmo.

Sobre esta matéria a União Europeia declarou: *“A digitalização é um passo essencial para gerar conteúdo digital que irá assegurar uma Europa digital. É uma actividade vital para preservar a herança cultural europeia, providenciando o acesso do cidadão a esse património, para melhorar a educação e o turismo, e para o desenvolvimento das indústrias de conteúdos digitais.”* (CALIMERA GUIDELINES, p.2)

De facto, a digitalização traz vários benefícios: o vasto e fácil acesso, a conservação de originais, a possibilidade de adicionar valor a imagens e colecções, e outras variadíssimas oportunidades como a divulgação de materiais e a atracção de um grande número de visitantes e utilizadores. A qualidade do ficheiro, o formato usado para o armazenar, a sua descrição, o seu uso, a sua preservação a longo prazo, o seu acesso aos utilizadores, a protecção contra infracções nos direitos de propriedade, são apenas algumas das questões que temos de considerar. Além disso, para cada tipo de material que é digitalizado há factores específicos que devem ser tidos em conta para garantir que o *output* digital possa ser utilizado com o máximo potencial, agora e no futuro.

Antes de iniciarmos a digitalização dos materiais é importante fazer uma avaliação do material a digitalizar, verificar o tamanho das obras, o seu formato e estado de conservação, no sentido de perceber se o equipamento de digitalização que dispomos é o adequado para o trabalho ou se é necessário adquirir outro.

Mas quais os requisitos técnicos que definimos como essenciais para realizar uma digitalização de qualidade? Segundo Lorna Hughs, “*A melhor abordagem é aquela que permite “digitalizar uma vez e para todos os propósitos” criando uma imagem matriz em alta qualidade da qual todas as outras versões (versões com compressão para disponibilizar via web) podem derivar. Esta master digital deve ser criada na mais alta resolução e profundidade de bit que é confortável e praticável, preservando na medida do possível a autenticidade e integridade da informação original.*” (HUGHES, 2004; 166 – 167). Esta abordagem é aconselhada por todos os teóricos destas matérias – criar uma imagem *master* da qual provirão todas as cópias. Como já referimos, a criação de uma verdadeira imagem *master* implica a pura representação do original, ou seja, tem de ter a máxima qualidade e estar isenta de qualquer tipo de compressão.

Numa perspectiva de gestão e de preservação a longo prazo é necessário definir princípios que resultarão na qualidade do material digitalizado, que têm a ver sobretudo com o manuseamento dos documentos, a utilização dos diferentes formatos e a denominação dos ficheiros:

Definição de normas técnicas para manuseamento

Os documentos, objectos a digitalizar, devem ser pré-classificados, de acordo com as suas características, de modo a organizá-los e agrupá-los. Esta classificação deve preceder o início do processo de digitalização.

As técnicas de manuseamento dos materiais a digitalizar podem ser variadas, dependendo do seu estado de conservação. As obras recentes e em bom estado de conservação, exigem apenas cuidados básicos, que se podem resumir a manter as mãos sempre limpas, evitando que a sujidade e a gordura presentes nas mãos sejam transpostas para o papel e ter um especial cuidado a remover fitas ou etiquetas de alarme presentes nos documentos. Quando se trata de obras antigas, estas necessitam de cuidados especiais de manuseamento, aconselhando-se ainda a utilização de luvas e máscaras.

Definição dos requisitos de qualidade para *masters* e cópias

É fundamental a definição prévia do que serão os requisitos de qualidade requeridos para a criação de ficheiros *masters* e cópias num processo de digitalização.

As imagens *master*, representações genuínas do original, serão a fonte para todas as cópias delas resultantes, ou seja, é a partir de uma imagem *master* que se fazem as cópias de menor resolução para fins de disponibilização na *Web*. Estas cópias, em JPEG, formatos mais leves e com alguma compressão, servem para fins de visualização e impressão enquanto a *master* serve para o propósito de arquivo. Desta forma, o utilizador só terá acesso às imagens em formatos JPEG porque é um formato mais leve e permite uma visualização ainda assim com qualidade. Outra vantagem da apresentação de cópias, relaciona-se com os direitos de publicação, trabalhar com baixas resoluções e qualidade de imagem implica que estas não possam ser usadas para publicação comercial ilegal.

Assim sendo, e segundo FERRER SAPENA [et al.] (2005), os requisitos para a constituição de um ficheiro de qualidade provêm da conjugação dos factores formato, resolução, resolução de cor e compressão, dependendo do tipo de documento que estamos a tratar. Por exemplo, no que respeita ao tipo de documento “*publicações periódicas*”, os parâmetros indicados são os seguintes:

Para a constituição de um ficheiro *master*:

TIPO DE DOCUMENTO	FORMATO	RESOLUÇÃO	PROFUNDIDADE DE COR	COMPRESSÃO
Publicações periódicas	TIFF	300 dpi	24 bits	Sem compressão

Para a constituição de um ficheiro que será disponibilizado via Internet – **cópias**:

TIPO DE DOCUMENTO	FORMATO	RESOLUÇÃO	PROFUNDIDADE DE COR	COMPRESSÃO
Publicações periódicas	JPEG, GIF, PDF	JPEG: 150 dpi	JPEG: 24 bits	JPEG: baixa
		GIF: 72 dpi	GIF: 8 bits	GIF: standard

FERRER SAPENA [et al.] (2005, p.105 e 106).

No que respeita ao PDF, é um formato de visualização que adopta as características (resolução, profundidade de cor, compressão) do formato original. Já o formato GIF (Graphics Interchange Format), apesar de mencionado no quadro, não é comumente utilizado para projectos desta natureza porque tem menos qualidade que o JPEG.

Definição de normas técnicas para a utilização de formatos

A definição de um formato é um factor preponderante num processo de digitalização. Um formato é aquilo que dá forma, ou seja, que estrutura um determinado conteúdo segundo tipologias documentais diferentes e objectivos diferentes. Deste modo, é

importante perceber quais os formatos que interessam à instituição, tendo em conta as suas necessidades.

Ao pensar num formato, temos de ter em conta a sua “usabilidade” e acessibilidade a longo prazo. Desta forma, provavelmente o uso de formatos standard e não proprietários (*software* livre) será a melhor solução a adoptar. Os formatos mais vulgarmente utilizados são os TIFF para as imagens de máxima resolução, “*masters*”, e o JPEG. Uma imagem digital pode ser gravada assim, em diferentes formatos.

O TIFF (*Tagged Image File Format*) é um formato utilizado para gravar “*masters*”, imagens digitalizadas com o propósito de arquivo e para efeitos de biblioteca digital. É um formato sem compressão, sem perda de qualidade, possuindo outra vantagem que é o facto de ser rico em meta-informação relativamente ao processo de produção da imagem, no entanto, muito pesado em termos de ocupação de espaço no disco.

Com base nas imagens *masters* são realizadas as cópias, em formato JPEG (*Joint Photographic Experts Group*), formato que já tem alguma compressão. Mas como se trata de um ficheiro de utilização, possibilita uma visualização e impressão de qualidade. É neste formato que a imagem é disponibilizada via *Web* para o utilizador, tanto em imagem como reprodução do original, como em *thumbnails* (ficheiros de imagem mais reduzidos), porque é um formato menos pesado a nível de disco.

Definição de normas técnicas para a denominação dos ficheiros

As imagens digitais ao serem gravadas necessitam de uma denominação. Uma colecção tem de ser sempre descrita no seu todo, no seu conjunto, deve ser única para todo o projecto e auto-explicativa uma vez que terá de permitir um cruzamento de referências de capítulo para capítulo, por exemplo, sem se perder. Neste sentido, e segundo o documento de “boas práticas” das bibliotecas da Universidade de Maryland (E.U.A), os nomes dos ficheiros devem seguir os seguintes parâmetros:

- a) *O nome do ficheiro tem de ser único e estruturado de forma consistente;*
- b) *Tem de levar em conta o número máximo de itens a serem digitalizados e direccionar isso para o número de dígitos a utilizar, se seguir um esquema numérico;*
- c) *Tem de conter zeros para facilitar a classificação em ordem numérica, se seguir um esquema numérico;*
- d) *Não pode conter um esquema excessivamente complexo ou longo de nomeação que pode ser susceptível a erro humano (...);*
- e) *Deve utilizar letras minúsculas e extensões de ficheiros;*

f) *Deve conter números e/ou letras, mas não caracteres como símbolos ou espaços que podem causar problemas através das plataformas operacionais.*

(CARIGNAN, 2007, p.12)

c) OS METADADOS

Os metadados são um conjunto de dados capazes de descrever outros dados, ou seja, é a informação estruturada agregada a um determinado dado que visa descrever as suas características para posteriormente facilitar a sua recuperação, descrição, identificação, uso, gestão e preservação.

Os metadados são uma parte chave da infra-estrutura informacional e são decisivos para ajudar a criar ordem no caos que é a Internet, é nesse sentido que descrevem, classificam e organizam a informação. Eles são basicamente linguagens de informação. Estas linguagens são compostas por: um vocabulário, conjunto de termos que a linguagem admite utilizar para definir os objectos que vai descrever, uma sintaxe, que é dada pelo formato de representação dos metadados (HTML, XML, RDF, por exemplo) e uma semântica, que dá o sentido aos termos e num segundo nível designa as relações que se estabelecem entre eles.

Todas as imagens ao serem digitalizadas trazem agregados metadados técnicos produzidos automaticamente pelas máquinas que as digitalizam, não sendo, no entanto, suficientes para descrever as imagens digitais, nem para garantir a sua preservação futura. Por esse motivo, é necessário atribuir também metadados descritivos e administrativos sobre a informação apresentada, estes têm de ser colocados “manualmente”. Os metadados administrativos respondem às necessidades de gestão dos próprios objectos e os descritivos, respondem às necessidades de recuperação dos mesmos.

A manutenção de uma biblioteca de objectos digitais exige a manutenção dos metadados sobre esses objectos. *“Sem metadados estruturais, os ficheiros com imagens ou texto serão de pouca utilidade, e sem metadados técnicos sobre o processo de digitalização, os investigadores poderão ter dúvidas sobre a exactidão da reflexão do original que a versão digital oferece. Por questões de gestão interna, a biblioteca deve ter ainda acesso a metadados técnicos apropriados para lhe permitir refrescar e migrar os dados, garantindo a durabilidade dos recursos.”* (BORBINHA, et al, 2007, p.2)

A NISO (2007) prenuncia-se também sobre os princípios que devem orientar a criação, gestão e utilização de metadados, de forma a garantir a sua qualidade:

1. Bons metadados correspondem a normas da comunidade no sentido em que são apropriados para os materiais da colecção, para os utilizadores e para a utilização corrente e futura da colecção;

2. Bons metadados suportam interoperabilidade;
3. Bons metadados usam controlo de autoridade e normas (standards) de conteúdo para descrever objectos;
4. Bons metadados incluem uma declaração clara das condições e termos de uso para os objectos digitais;
5. Bons metadados suportam a preservação a longo-prazo dos objectos nas colecções.

Os ficheiros de metadados devem conter a qualidade de bons objectos digitais, incluindo a autoridade, autenticidade, capacidade de arquivo, persistência e um identificador único.

d) O CONTROLO DE QUALIDADE

O controlo de qualidade é a implementação de processos e técnicas no sentido de verificar a qualidade, precisão e consistência do material digital. Após a selecção do material, a sua digitalização e posterior inserção de metadados é necessário fazer um controlo de qualidade de todas as imagens digitalizadas, comparando-as com o original e verificando se existem erros a assinalar.

O documento *Digital Imaging for Photographic Collections: Foundations for Technical Standards* aborda esta questão. A melhor abordagem para o controlo de qualidade de imagem digital implica, por um lado, a inspecção visual subjectiva e, por outro lado, medidas de inspecção objectivas realizadas através do *software* nos próprios ficheiros digitais. (FREY; REILLY, 2006)

Existem directrizes que podem ser utilizadas como suporte para efectuar um controlo de qualidade eficiente e eficaz. Exemplo disso é o documento “*Features: Practical quality control procedures for digital imaging projects*” desenvolvido pelo Digital Media and Image Center (DMIC) através de um programa da Biblioteca Digital da Universidade de Indiana (E.U.A) que apresenta uma metodologia para a realização do controlo de qualidade.

Para esta instituição, o controlo de qualidade divide-se em dois processos. O processo de “controlo automático”, que verifica basicamente se o formato do arquivo, o sistema de compressão, a resolução de imagem, entre outros, está de acordo com especificações do projecto, e ainda o processo de “controlo visual”, que implica uma verificação manual das imagens. Este processo tenta verificar se o nome do arquivo atribuído à imagem está correcto com base no seu número de identificação, se a cor está igual ao original, se a imagem digital está com a orientação correcta, se existem partículas

de papel, pó ou pêlos nas imagens, se as imagens aparecem pela ordem correcta, se os metadados estão correctos, se não existe omissão de texto ou números de página, etc.

Caso se encontre alguma anomalia na colecção digitalizada, essa falha é registada e posteriormente corrigida. Através deste controlo rigoroso a instituição garante uma colecção digital de qualidade que respeita todos os parâmetros pré-definidos.

e) A DISPONIBILIZAÇÃO

O processo de disponibilização é a etapa que se segue após terminado o controlo de qualidade e as possíveis correcções. Esta etapa é a “última” fase de constituição efectiva de um projecto de biblioteca digital, pois é a responsável pela apresentação do resultado final de todo o processo.

Quando falamos em disponibilização referimo-nos ao processo de tornar os conteúdos digitais acessíveis através da Internet. Existem vários programas que permitem fazê-lo, tanto em formato proprietário como em *software* livre.

Dado que se tratam de conteúdos digitais que terão de ser geridos e mantidos por longos períodos de tempo, provavelmente a melhor estratégia passará pela utilização de programas não-proprietários, programas que podem facilmente ser alterados e actualizados segundo as necessidades da instituição que os acolhe. “*A melhor forma de assegurar um acesso a longo termo a dados digitais é usando formatos standards e sistemas de código aberto (como OAIS - Open Archive Information System) sempre que possível*” (HUGHES, 2004, p. 205)

Neste campo, a instituição terá de decidir se disponibilizará estes conteúdos através da Internet, mundialmente, sem restrições, ou se haverá a necessidade de restringir o acesso aos conteúdos por algum motivo (constrangimentos de direitos de autor, por exemplo).

f) MANUTENÇÃO DO PROJECTO

O processo de implementação de um projecto de biblioteca digital não acaba com a sua disponibilização *On-line*, isto porque a gestão dessas colecções digitais é um processo contínuo.

Os objectos digitais têm um “ciclo de vida”, ou seja, ao longo da sua existência passam por várias fases que normalmente se traduzem no nascimento, crescimento, maturidade e possível declínio do objecto digital. Logo, existe a necessidade de planear não só a criação dos ficheiros digitais, como também a sua gestão ao longo do tempo, e considerar a possível conversão em outro formato para ser re-utilizado (migração).

O “ciclo de vida” de um objecto digital é contínuo e a sua manutenção tem de ser cuidada, pois caso contrário, pode originar obsolescência dos conteúdos digitais. Neste campo, a importância da existência dos metadados é preponderante, os metadados tem de percorrer todas as fases do ciclo de vida dos recursos digitais para que estes sejam recuperados e preservados ao longo do tempo.

O conceito de preservação digital pode definir-se segundo Miguel Ferreira, como a actividade responsável por garantir que a “*comunicação entre um emissor e um receptor é possível, não só através do espaço mas também através do tempo*” (2006, p. 12). No fundo, a preservação digital consiste na capacidade de garantir que a informação digital permaneça acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação.

CAP. III – BIBLIOTECA DIGITAL DA FLUL (2ª FASE). O PROJECTO DE DISPONIBILIZAÇÃO DE PUBLICAÇÕES PERIÓDICAS

3.1 - CARACTERIZAÇÃO DA INSTITUIÇÃO QUE ACOLHE O PROJECTO

3.1.1 - A BIBLIOTECA DA FACULDADE DE LETRAS

A biblioteca da Faculdade de Letras da Universidade de Lisboa é originária da Biblioteca do Curso Superior de Letras² e as suas raízes remontam ao século XIX.

Os seus depósitos contêm actualmente um acervo rico em obras de várias áreas culturais que vão desde o século XV à actualidade e foi constituído através da aquisição de obras, por via de doações como as da Academia Real das Ciências de Lisboa e da Academia de História, e ainda, pela oferta de professores da Faculdade e de intelectuais portugueses.

A FLUL foi criada, formalmente, em 1911 e só no final dos anos 50, a biblioteca, juntamente com a Faculdade, foram transferidas do edifício anexo à Academia das Ciências para as instalações da Cidade Universitária. A biblioteca continuou a prestar serviços nestas instalações nos mesmos moldes que prestava anteriormente, sem sofrer grandes transformações até ao final dos anos 80, período em que se procedeu a uma modernização e reorganização da biblioteca com a implementação de um processo de informatização, naturalmente necessário e premente.

Em termos logísticos, a biblioteca sofreu uma alteração muito positiva em 2000. O espólio que antes se encontrava espalhado por várias bibliotecas na faculdade foi reunido num edifício novo projectado pelo arquitecto holandês Harro Wittmer com uma área coberta de 6.600 m². Conjuntamente a este edifício, construiu-se uma sala de conferências e reuniões, um anfiteatro e espaços onde se prestam serviços de apoio. À modernidade e funcionalidade do edifício aliou-se a reestruturação das funções da biblioteca com a disponibilização de um variado leque de novos serviços.

Presentemente, a biblioteca da FLUL possui cerca de 300.000 monografias e mais de 1600 títulos de publicações periódicas e disponibiliza serviços que visam responder às

² O Curso Superior de Letras em Lisboa foi fundado em 1859 pelo rei D. Pedro V. Este curso constituiu o núcleo de estudos superiores de humanidades em Portugal sendo o antecedente da actual Faculdade de Letras da Universidade de Lisboa.

necessidades dos utilizadores directos daquela biblioteca – os estudantes da Faculdade de Letras, mas também ao crescente número de estudantes de outras faculdades que escolhem esta biblioteca como local de estudo. A Biblioteca é uma das bibliotecas cooperantes do SIBUL (Sistema Integrado de Bibliotecas da Universidade de Lisboa) e participa na manutenção e actualização deste catálogo colectivo.

No entanto, o que define uma organização e lhe dá identidade é a sua missão. A missão da biblioteca da FLUL consiste em *“Adquirir, processar, conservar e tornar acessíveis os recursos de informação, de tipologias e suportes diversificados, especializados nas áreas de ensino e investigação da FLUL.”* (Biblioteca da FLUL).

3.1.1.1- A BIBLIOTECA DIGITAL DA FLUL

A Biblioteca Digital da FLUL surgiu em 2005 com o apoio do Programa Operacional de Cultura (POC) promovido pelo Ministério da Cultura. A instituição, que procurava outros meios de disponibilização da informação e atingir novos públicos, candidatou-se a este programa com o intuito de conseguir verbas para o desenvolvimento de um projecto de biblioteca digital que previa a conservação, digitalização e difusão do património bibliográfico e documental da FLUL. Ao ser aprovado este programa, disponibilizou os meios financeiros necessários para a sua implementação.

A missão da biblioteca digital da FLUL é clara: *“Integrada na estratégia de desenvolvimento e modernização da Biblioteca da FLUL, a Biblioteca Digital, (...) cumpre, apoiando-se nas mais modernas e inovadoras tecnologias de informação e comunicação, a missão de conservar e tornar acessíveis estes recursos de informação a todos quantos deles queiram usufruir, nomeadamente aos investigadores, garantindo-lhes um acesso e visibilidade universais.* (Biblioteca digital da FLUL).

A biblioteca digital começou por disponibilizar *On-line* duas colecções distintas, a colecção de incunábulos e impressos raros (séculos XV e XVI) e a colecção de obras de teatro do espólio de Osório Mateus. Tendo em mente o sucesso deste primeiro projecto e procurando disponibilizar novos conteúdos *On-line*, a biblioteca digital apresentou, em 2008, uma 2ª fase do projecto de biblioteca digital totalmente suportado pela instituição - *O Projecto de disponibilização de publicações periódicas editados pelos Departamentos e Centros de Investigação da Faculdade de Letras da Universidade de Lisboa.*

3.2 – O PROJECTO

O projecto que apresentaremos de seguida constitui a 2ª fase do projecto Biblioteca Digital desenvolvido pela biblioteca da Faculdade de Letras da Universidade de Lisboa. Este projecto direcciona-se para as *Publicações periódicas* editadas pelos Departamentos e Centros de Investigação da Faculdade e pretende disponibilizar estas revistas via biblioteca digital. Deste modo, visa dar continuidade ao projecto base da Biblioteca Digital da FLUL, contribuindo com novos conteúdos *On-line*. (ANEXO 1 – Cronograma da constituição do projecto).

Depois da apresentação de algumas características base que distinguem este projecto, segue-se a definição do seu público-alvo assim como os motivos pelos quais se optou por seleccionar estas obras em particular. Posteriormente, identificaremos o problema que o projecto enfrenta e vem responder, faremos uma análise dos pontos fortes e fracos do projecto e procedemos à definição de objectivos, para finalizar, faremos uma demonstração dos diferentes processos que levam à conclusão do mesmo.

CARACTERÍSTICAS BASE:

A Biblioteca Digital da FLUL iniciou-se com um projecto realizado em *Outsourcing*, ou seja, feito por entidades externas que possuíam recursos humanos e tecnológicos próprios e efectuaram a digitalização dos materiais, preservação e a respectiva disponibilização em linha³. No entanto, nesta 2ª fase, do projecto de biblioteca digital optou-se por desenvolver um projecto totalmente realizado *In-house*. Porquê? No projecto anterior (1ª fase), as colecções a digitalizar precisavam de cuidados e de tratamentos específicos. As obras da colecção de incunábulos e impressos raros dos séculos XV e XVI tinham de ser manuseadas com extremo cuidado e a digitalização das suas páginas só era possível através de um equipamento devidamente apropriado para o efeito. O mesmo sucedeu, de certa forma, com a colecção de obras de teatro do espólio de Osório Mateus, exemplares únicos que teriam de ser digitalizados em *scanners* apropriados. No projecto actual, a exigência a este nível é menor devido à condição razoável dos exemplares objectos de digitalização, o que possibilitou que se prescindisse da necessidade acrescida de preservação ou extremo cuidado no manuseamento.

Outro factor preponderante e que levou a que se desenvolvesse o projecto *In-house* foi o factor monetário. O projecto anterior contava com um apoio financeiro (POC) que

³ A digitalização, nomeação e metadados foi desenvolvida pela empresa *Scansystem*. A indexação foi feita pela empresa *DigiCult*.

suportava todos os custos inerentes a um projecto desta envergadura. Sem esses meios financeiros a única opção seria digitalizar estas publicações nas instalações da biblioteca, utilizando os seus equipamentos, os recursos humanos existentes e o trabalho desenvolvido por voluntários.

Os recursos humanos direccionados para o projecto foram constituídos por um funcionário da biblioteca, que não estando em permanência no projecto, orientava os trabalhos⁴; quatro voluntários, que despendiam de algumas horas do seu tempo para a instituição, em diferentes períodos; uma estagiária em regime pós-laboral e ainda pelo Coordenador do Projecto, o Chefe de Divisão da Biblioteca, Dr. Pedro Estácio.

3.2.1 – PÚBLICO-ALVO

O público-alvo deste projecto é formado por estudantes e investigadores, sendo de acesso aberto, aos cidadãos em geral. No entanto, considerámos como público-alvo primário a comunidade científica nacional e internacional, nomeadamente a dos países de língua portuguesa.

3.2.2 - SELECÇÃO DA COLECÇÃO

A biblioteca digital da FLUL considerou pertinente e desejável enriquecer a já existente biblioteca digital com a disponibilização *On-line* destas publicações em particular. Perante esta opção coloca-se a seguinte questão: Porquê digitalizar estas publicações periódicas e não outras colecções? A resposta é-nos dada em parte pelo documento “*A Biblioteca Digital da Faculdade de Letras. Em busca da trivialidade*” de um dos coordenadores da Biblioteca Digital da FLUL, Professor Doutor João Dionísio. De todo o espólio existente na biblioteca foi pertinente digitalizar este conjunto de colecções pelas razões enumeradas no documento:

- I. *O interesse institucional de tornar acessível a produção científica periódica da Faculdade;*
- II. *A continuação de acções de digitalização já desenvolvidas por algumas das revistas (Finisterra, Boletim de Filologia, Revista Lusitana),*
- III. *Em termos de dimensão, a constituição de um corpus por um lado significativo, por outro trabalhável num prazo razoável,*

⁴ Inicialmente com a contribuição da Sr.^a Dr.^a Cristina Faria, e posteriormente com a orientação da Sr.^a D.^a Graça Lopes.

- IV. *A concentração em bibliografia portuguesa difícil de pesquisar (em condições normais, os únicos artigos destes periódicos pesquisáveis em catálogo são os que foram publicados também sob a forma de separata),*
- V. *A possibilidade de por esta via serem gerados núcleos de bibliografia crítica interessante para ser articulada com outras acções e projectos, designadamente com as colecções do livro antigo e do espólio Osório Mateus.*

(DIONISIO, p. 7 e 8)

Como mencionámos no ponto 2.4 do Cap. II, os critérios de selecção do material a digitalizar podem ser de diversa ordem. Neste caso concreto, podemos concluir que as razões pelas quais estas revistas foram escolhidas se deveram a factores como os apontados, por exemplo, pelo documento *Selection Criteria for Preservation Digital Reformatting* (2002) da Biblioteca do Congresso: valor, uso, condição e características do original.

Neste sentido, o parâmetro *valor* relaciona-se com o “interesse institucional”, *uso* – “bibliografia de outra forma difícil de pesquisar”, *condição* – “algumas revistas já tinham iniciado processos de digitalização” e *características do original* – sem necessidades de cuidados especiais e cujo conteúdo literário se destaca por se tratar de produção científica nacional.

A instituição elegeu estas revistas porque possui as suas séries completas e a maior parte delas tem poucos números publicados, o que possibilita um trabalho célere. Se no projecto anterior a colecção era patrimonial e o seu fundamento era dar acesso e preservar, nesta o fundamento passa por apresentar *On-line* colecções mais recentes que por não precisarem de cuidados especiais na digitalização, nem de intervenções a nível da preservação, necessitam de meios menos exigentes. Todos estes factores conjugados tornam o processo de digitalização mais simples ao nível financeiro, humano e técnico.

Como referimos, o objecto deste projecto são publicações periódicas. Contudo, como podemos definir um periódico? Um periódico é uma publicação em série, “*Considera-se publicação em série uma obra, editada em partes sucessivas, tendo usualmente designação numérica e/ou cronológica, e concebida para ser continuada por período indefinido*”. (Regras portuguesas de catalogação, 2000, p. 61). Pode tratar de vários assuntos ou de um assunto específico, a sua periodicidade pode ser regular (diário, semanal, quinzenal, mensal, bimestral, trimestral, quadrimestral, semestral, anual,...) ou irregular sem intervalos certos. Uma particularidade fundamental deste tipo de publicação é a continuidade, ou seja, a publicação dos números tem uma duração indeterminada. Cada publicação é um número e um conjunto de determinados números estabelece um volume.

3.2.3 - IDENTIFICAÇÃO DO PROBLEMA

Tendo em conta o sucesso das primeiras colecções de livro antigo disponibilizadas em linha, impõe-se uma questão. Porquê transitar para uma 2ª fase da biblioteca digital e não dar continuidade ao primeiro projecto? A biblioteca digital da FLUL, e mais especificamente os seus coordenadores, Dr. Pedro Estácio, Chefe de Divisão da Biblioteca da Faculdade de Letras e o Sr. Doutor João Dionísio, professor da Faculdade, optaram por não ampliar a colecção já existente com outros livros do fundo antigo, mas incluir outros objectos também importantes e que constituíam uma opção claramente mais económica.

O projecto de disponibilização em linha de publicações periódicas apresenta algumas particularidades, que o diferenciam indubitavelmente do anterior. São estas diferenças que o caracterizam e singularizam. Podemos começar por enunciar a divergência de conteúdos: o primeiro destacando-se com a apresentação de conteúdos patrimoniais, este disponibilizando revistas científicas actuais; o primeiro sem o constrangimento dos direitos de autor, o segundo obrigado a enfrentar esta realidade. Ambos dispõem de recursos humanos, financeiros e técnicos distintos, o primeiro desenvolvido em *Outsourcing*, este *In-house*.

Como implementar um projecto desta natureza, com estas características, numa biblioteca digital, tendo como antecessor um projecto já com provas dadas de qualidade e sucesso? De forma sucinta, poderemos afirmar que tentaremos manter os padrões de qualidade existentes, adaptando as soluções às características deste projecto em particular.

Na prática, um problema mais específico se adivinha: como digitalizar as imagens, organizá-las de forma coerente e dotá-las de metadados, de forma a estruturá-las, permitindo a sua reutilização em qualquer contexto agora e no futuro? Por fim, como disponibilizar este conteúdo *On-line*? São estas as questões que se pretendem responder no desenvolvimento deste trabalho.

No entanto, para podermos intervir numa determinada realidade necessitamos de fazer um estudo prévio que nos permita verificar os aspectos positivos e negativos nela envolvidos. Será importante realizar um diagnóstico da situação daquele universo em particular.

3.2.4 – AVALIAÇÃO OU DIAGNÓSTICO DA SITUAÇÃO – ANÁLISE SWOT

A análise SWOT é uma “ferramenta” usada para efectuar planeamento estratégico através de uma análise de cenário da realidade da instituição. A sigla *SWOT* corresponde às

designações *Strengths* (forças), *Weaknesses* (fraquezas), *Opportunities* (oportunidades) e *Threats* (ameaças) de uma determinada organização ou sistema e ajuda a identificar as características internas do projecto e a forma como se relaciona com o meio envolvente, o exterior. Optámos por utilizar esta ferramenta de diagnóstico porque considerámos ser um método eficiente para analisar as características principais deste projecto.

MATRIZ SWOT – Análise do projecto de digitalização dos periódicos da FLUL

	FORÇAS (Ajuda a atingir o objectivo)	FRAQUEZAS (Ameaça o atingir do objectivo)
Interna (Atributos do da Organização)	<ul style="list-style-type: none"> ▪ Projecto <i>In-house</i> ▪ Obras de estrutura linear e com muitos exemplares ▪ Experiência do projecto anterior ▪ Adopção de normas BND ▪ Recursos técnicos – equipamento e programas (<i>Hardware e Software</i>) adequados ao projecto de digitalização ▪ Experiência dos funcionários ▪ Rigor na digitalização, implementação de metadados, controlo de qualidade 	<ul style="list-style-type: none"> ▪ Direitos de autor (Copyright) ▪ Planeamento ▪ Recursos humanos – Voluntários sem especialização na área de biblioteconomia ou arquivística ▪ Recursos financeiros ▪ Condicionantes do programa ContentE ▪ Ausência de <i>know-how</i> técnico na instituição sobre a tecnologia associada (<i>software</i> livre)
	OPORTUNIDADES	AMEAÇAS
Externa (Atributos do da Ambiente)	<ul style="list-style-type: none"> ▪ Actualizações em desenvolvimento nos programas PAPAIA e ContentE ▪ Possibilidade de estabelecer acordos de cooperação com outras bibliotecas digitais 	<ul style="list-style-type: none"> ▪ Obsolescência da tecnologia

Tabela 1 – Matriz de análise SWOT

ANÁLISE INTERNA – FORÇAS e FRAQUEZAS

A análise interna proporciona a identificação e monitorização do ambiente interno da organização, neste caso, permite-nos verificar quais as características que fortalecem ou enfraquecem este projecto.

FORÇAS

1º PONTO – PROJECTO *IN-HOUSE*

Apontámos como aspecto positivo o facto deste projecto se desenvolver totalmente *In-house*. O projecto constitui, neste sentido, uma inovação pois tem a incumbência de se fazer valer dos conhecimentos, recursos humanos e financeiros da instituição, implementando um projecto totalmente desenvolvido nas instalações da Faculdade de Letras (Biblioteca e reprografias).

O conhecimento adquirido com o desenvolvimento de um projecto de biblioteca digital *In-house* capacita a biblioteca de um *Know-how* extremamente importante a ser canalizado para projectos futuros. Este *Know-how* está relacionado com a experiência adquirida com a realização de todos os processos que levam à constituição de uma biblioteca digital. Além do conhecimento adquirido que poderá ser aplicável no futuro, a biblioteca possuirá funcionários especializados nesta área em particular, o que será uma mais valia a registar.

2º PONTO – OBRAS DE ESTRUTURA LINEAR E COM MUITOS EXEMPLARES

A tipologia dos exemplares e o facto de existirem em grande número, constituem um ponto forte neste projecto. O formato linear das obras permite uma desenvoltura no manuseamento dos exemplares e poucas alterações no processo de digitalização, já o facto de a maioria das colecções de periódicos possuírem vários exemplares de cada número, permite que possamos usar um exemplar para ser guilhotinado e ser facilmente colocado num alimentador de *scanner* para ser digitalizado automaticamente.

3º PONTO – EXPERIÊNCIA DO PROJECTO ANTERIOR

A experiência resultante do projecto anteriormente desenvolvido na biblioteca digital é uma mais valia a salientar. Apesar de todas as diferenças de base existentes entre os projectos, a experiência anterior vem dotar este projecto de conhecimentos empíricos importantíssimos na gestão e execução de todos os processos envolvidos. Desta forma, este projecto acaba por seguir os passos do anterior em toda a sua estrutura, isto é, segue-o como modelo e referência.

4º PONTO – ADOPÇÃO DE NORMAS BND

Tanto o projecto anterior como este visa seguir tanto quanto possível as práticas e normas utilizadas pela BND, modelo de referência da biblioteca na área das bibliotecas digitais. A manutenção de uma biblioteca digital exige a conservação dos metadados sobre os objectos digitais, para resolver esta questão a BN adoptou o esquema METS (Metadata

Encoding and Transmission Standard)⁵ que se expressa usando a linguagem XML (eXtensible Markup Language)⁶, esquemas utilizados também neste projecto.

5º PONTO – RECURSOS TÉCNICOS

A instituição possui os requisitos técnicos, tanto a nível de equipamentos *hardware* como de *software*, necessários para a execução das tarefas definidas. Os trabalhos decorrem nas instalações do arquivo da biblioteca da FLUL, habilitadas com dois postos de trabalho, sendo que cada um é constituído por um monitor e um computador. Para fazer face às necessidades de um projecto desta natureza a biblioteca adquiriu dois *scanners*, um *scanner* de mesa ou *flatbed* (A3), e um *scanner* com alimentador, ambos da marca *Scanner Kodak i1220*.

6º PONTO – EXPERIÊNCIA DOS FUNCIONÁRIOS

Neste campo, destacamos como factor positivo a experiência dos funcionários da biblioteca detentores de conhecimentos que advinham da prática no projecto anterior e ainda o trabalho e empenho da estagiária e de uma voluntária especializadas na área de biblioteconomia.

7º PONTO – RIGOR

De facto, este projecto seguia directrizes rigorosas no que respeita aos processos de digitalização, implementação de metadados e controlo de qualidade. Estes processos foram realizados segundo critérios normalizados e de acordo com as exigências biblioteconómicas.

FRAQUEZAS

1º PONTO – DIREITOS DE AUTOR

Um aspecto importante que poderá dificultar o desenvolvimento deste projecto e que se apresenta deste modo como um ponto fraco é o direito de propriedade sobre os materiais a trabalhar. É necessário obter permissão dos autores, no âmbito dos direitos de

⁵ O METS é um formato *standard* para a codificação e transmissão de metadados, pensado principalmente para o envio de ficheiros, imagens e objectos multimédia de uma biblioteca digital. O METS é uma iniciativa da Digital Library Federation e para a utilização do METS usa-se a estrutura de etiquetas XML.

⁶ O XML proporciona um armazenamento de dados estável e não proprietário bem com um ambiente de troca no qual os dados podem ser transferidos para outros ambientes de software se necessário.

autor para digitalizar as colecções seleccionadas, e posteriormente obter autorização para as disponibilizar.

A Faculdade de Letras é a proprietária das publicações, facto que poderá tornar a questão dos direitos de autor mais fácil de resolver, em virtude de existir uma relação directa e de proximidade entre a instituição e os autores desses artigos. Contudo, esta questão não é linear, como tal a biblioteca terá de tomar medidas no sentido de ultrapassar este problema.

2º PONTO – PLANEAMENTO

A biblioteca optou por não desenvolver um planeamento rigoroso das várias fases de implementação do projecto porque decidiu fazer uma planificação segundo a evolução dos trabalhos. Apontámos este ponto como uma fraqueza. Se por um lado percebemos esta decisão, não havia custos financeiros ou de mão-de-obra envolvidos que obrigassem a uma planificação rigorosa, por outro, consideramos que seria mais positivo para o projecto se existisse um planeamento efectivo, com prazos, com metas, que atribuísem aos funcionários e voluntários responsabilidades pelo atraso ou não do projecto.

Este tipo de planeamento, feito segundo a evolução dos trabalhos, como que consente a existência de um sentimento de “é para se ir fazendo”, contudo, com um planeamento efectivo, muito provavelmente, as pessoas envolvidas no projecto sentir-se-iam no dever de concretizar o projecto nos prazos determinados.

3º PONTO – RECURSOS HUMANOS – VOLUNTÁRIOS SEM ESPECIALIZAÇÃO NA ÁREA DE BIBLIOTECONOMIA OU ARQUIVÍSTICA.

Registámos como ponto fraco, a presença sazonal de voluntários cuja ausência de formação na área de biblioteconomia e arquivística, aliada à fraca assiduidade, obrigaria à aplicação de formações contínuas que se revelavam pouco produtivas.

4º PONTO – RECURSOS FINANCEIROS

O projecto não possuía meios financeiros para a execução das suas tarefas. De facto, só se registavam custos associados à compra de equipamentos, postos de trabalho e *scanners*, e estes custos eram suportados pelo orçamento da Faculdade. No que respeita aos custos de mão-de-obra eram inexistentes. A mão-de-obra era voluntária e não existia nenhum funcionário da instituição em permanência no projecto.

O facto de não existirem recursos financeiros direccionados para o projecto é um ponto fraco a registar. Consideramos que seria de grande importância a direcção de custos

para a contratação de um técnico especializado na área e a sua presença a tempo inteiro no projecto.

5º PONTO – CONDICIONANTES DO PROGRAMA CONTENTE

A finalização do projecto e a consequente disponibilização *On-line* dependem do programa ContentE. Contudo, este programa apresenta duas condicionantes que podem pôr em causa a continuidade do projecto. A primeira condicionante prende-se com a necessidade de fazer alterações no próprio *software*. A biblioteca pretende ter uma versão personalizada do *software*, que contenha as características específicas da Biblioteca da Faculdade de Letras actualizadas, particularmente, no que respeita às “folha de estilos”. Se esta alteração não for conseguida atempadamente pode colocar em causa a disponibilização das revistas já digitalizadas.

A segunda condicionante prende-se com a funcionalidade do programa. Consideramos o programa pouco intuitivo e de difícil manuseamento e utilização.

6º PONTO – AUSÊNCIA DE *KNOW-HOW* TÉCNICO NA INSTITUIÇÃO SOBRE A TECNOLOGIA ASSOCIADA (*SOFTWARE* LIVRE)

A instituição optou por utilizar “*software* livre” no desenvolvimento deste projecto que consiste nos programas PAPAIA e ContentE, criados, desenvolvidos e utilizados pela Biblioteca Nacional Digital (BND) e com provas dadas de eficiência no projecto de biblioteca digital anterior. Considerámos que a aposta em *softwares* de código aberto é uma valência para a realização do projecto e principalmente para a sua continuidade.

O *software* livre é muito mais do que um *software* gratuito. O facto de ser livre significa que seu código-fonte é passível de ser alterado de modo a adequá-lo às necessidades da instituição ou da pessoa que o utiliza, prescindindo de qualquer investimento financeiro. Logo, revela-se uma mais valia mesmo em termos de preservação.

As principais vantagens deste tipo de programas são as seguintes: evoluem segundo os interesses da biblioteca ou comunidade que o utiliza e não segundo os interesses da empresa criadora; permitem a modificação do *software* para uma funcionalidade específica; reduzem os custos em toda a cadeia de uso do *software*; constituem aplicações informáticas normalizadas e sustentáveis a longo prazo.

As desvantagens existem, mas são pouco relevantes. Podem resumir-se a dificuldades de adaptação por parte dos utilizadores habituados às plataformas *Windows* e ao facto de se correr o risco, com este tipo de programas, de se tornarem obsoletos se a comunidade não se interessar pelo desenvolvimento dos mesmos.

Contudo, neste projecto, a utilização de *software* livre pode ser contraprodutiva, pois não existem na instituição técnicos capazes de adequar o *software* às necessidades da biblioteca. Consideramos que um *software* livre só é uma mais valia para um projecto se existirem técnicos com o *know-how* necessário para procederem à sua actualização, por exemplo.

Tendo em conta este diagnóstico, no que respeita às fraquezas existentes neste projecto, podemos retirar algumas conclusões. A continuidade e consequente concretização deste projecto encontram-se ameaçadas essencialmente por dois factores, ambos de ordem interna. O primeiro devido à opção da instituição em utilizar o *software* ContentE, mesmo sabendo que este necessitava ainda de ser personalizado, e o segundo, devido à opção da biblioteca em utilizar recursos humanos em regime de voluntariado para desenvolver o projecto.

A adequação do programa ContentE às necessidades da biblioteca é preponderante para a continuidade do projecto. Esta personalização do programa à biblioteca digital da FLUL foi solicitada à BN, mas de facto, enquanto a biblioteca não conseguir a especificação do programa ContentE, isto é, a concretização das alterações ao *layout* de apresentação *On-line* já predefinido, será impossível disponibilizar os conteúdos já digitalizados no sítio da biblioteca digital, ou seja, o projecto não será finalizado, o que resultará na sua indisponibilidade para o exterior.

Outra ameaça à continuidade do projecto prende-se com os recursos humanos nele envolvidos. O trabalho voluntário, cuja área de formação na sua maioria não se relaciona com a biblioteconomia ou arquivística, pode revelar-se contraproducente. Em termos gerais, é possível que não exista o retorno esperado das formações que foram prestadas devido à pouca assiduidade e interesse dos mesmos. Estes factores, aliado ao facto de não existir um funcionário da biblioteca a tempo inteiro no projecto, levará a que seja muito difícil a concretização do mesmo, a sua conclusão e consequente disponibilização para o exterior.

Análise externa – OPORTUNIDADES e AMEAÇAS

A análise externa permite a identificação das principais oportunidades e ameaças que pairam sobre a organização/instituição. Esta análise prevê os desenvolvimentos futuros que possam ter maior ou menor impacto na organização e que estão totalmente fora de controlo da mesma.

OPORTUNIDADES

1º PONTO – ACTUALIZAÇÕES DOS PROGRAMAS PAPAIA E CONTENTE

Os *softwares* PAPAIA e ContentE encontram-se de momento em fase de actualização e aperfeiçoamento por parte da BND. As versões actualizadas destes programas resultarão na simplificação do trabalho de implementação de metadados e estruturação de conteúdos, facilitando o surgimento de novos projectos e o consequente crescimento da biblioteca digital da FLUL

2º PONTO – POSSIBILIDADE DE ESTABELECEMOS ACORDOS DE COOPERAÇÃO COM OUTRAS BIBLIOTECAS DIGITAIS

A biblioteca digital da FLUL mantém um acordo de cooperação com a BND para a qual disponibiliza obras de livro antigo digitalizadas ao abrigo do anterior projecto de biblioteca digital.

O actual projecto de digitalização de publicações periódicas científicas pode originar a proliferação de acordos de cooperação com outras bibliotecas nacionais e internacionais que serão úteis para a troca de informação e experiências.

AMEAÇAS

1º PONTO – OBSOLESCÊNCIA TECNOLÓGICA

A ameaça externa que pode estar associada a um projecto desta natureza relaciona-se com o problema da preservação digital e a inevitável obsolescência da tecnologia.

A preservação não é linear. O suporte físico onde a informação é armazenada está em constante mutação, o formato em que é armazenada pode não ser compatível com futuras máquinas, o *hardware* e *software* estão em constante actualização. Assim, a problemática da preservação digital surge porque o documento electrónico (nativo ou procedente da digitalização) está dependente de um sistema intermediário para a decodificação e interpretação do seu conteúdo.

De facto, o material digital transporta consigo um problema estrutural que coloca em risco a sua longevidade porque, embora um documento digital possa ser copiado infinitamente sem perda de qualidade, este documento necessita sempre de um contexto tecnológico que o suporte. Esta dependência tecnológica torna-o assim vulnerável à rápida obsolescência dos suportes que o armazenam.

3.2.5 – OBJECTIVOS E PLANEAMENTO

A 2ª fase da biblioteca digital procura “atingir outros objectos distintos dos do Projecto Bibliográfico e Documental, alargar públicos, concentrar recursos, articulá-los em colaboração sempre que possível.” (DIONISIO, 2008, p. 6). Neste sentido, o projecto de Publicações Periódicas surgiu com o objectivo de “diversificar a oferta, atingir novos públicos, centralizar e articular recursos” (DIONISIO, 2008, p. 9).

Podemos afirmar que este Projecto almeja um objectivo de carácter geral que consiste em disponibilizar estas publicações via Internet, e uma meta de carácter específico, que seria o de digitalizar seis títulos dessas revistas até ao final do ano de 2009.

Como foi referido na análise SWOT efectuada, neste projecto não existiu um planeamento rigoroso dos trabalhos, porque a própria natureza do projecto não o exigia (pois não tinha custos financeiros ou humanos associados). Os responsáveis por este projecto consideraram que a melhor opção consistia no planeamento gradual das tarefas, baseado no seu ritmo de desenvolvimento, em virtude de ser contraproducente planeá-lo de forma rigorosa, fixar prazos rígidos e definir metas muito precisas quando o projecto dependia em grande parte do trabalho sazonal de voluntários.

3.2.6 - IMPLEMENTAÇÃO DO PROJECTO

Antes de se proceder à implementação do projecto, a biblioteca tomou decisões que possibilitariam a resolução dos problemas relativos aos direitos de autor. Como mencionámos no ponto 3.2.4, na análise SWOT, esta era uma questão decisiva que tinha de ser solucionada.

RESOLUÇÕES – DIREITOS DE AUTOR

O direito de autor visa proteger obras onde os seus autores exprimem as suas ideias, ideologias, pensamentos e filosofias de vida. As criações desses autores podem ser do domínio literário, científico ou artístico. O período legal de protecção dos seus direitos corresponde à vida do autor e aos 70 anos subsequentes à sua morte. Deste modo, é indispensável que se analise quais os direitos de propriedade existentes sobre o material a digitalizar e só após esta definição, delinear as formas de desenvolver o nosso trabalho, respeitando e obtendo a anuência dos seus autores.

Neste projecto, a resolução desta questão era fundamental, pois dela dependia o futuro do mesmo. Trata-se da disponibilização em linha de publicações periódicas científicas actuais que, além de deterem direitos de autor associados, podem ainda pôr em

causa a comercialização das mesmas revistas em suporte analógico. Tendo em conta esta problemática, como resolver a questão relativa à disponibilização em acesso livre destas revistas? Neste caso, a biblioteca determinou digitalizar todas as revistas mantendo as suas imagens em arquivo até haver efectiva autorização para serem disponibilizadas *On-line*.

A biblioteca optou por uma estratégia que se delineia através de várias acções com o objectivo de ultrapassar este dilema:

- a) Antes de se proceder à disponibilização dos conteúdos digitalizados, serão contactados os editores das revistas e respectivos autores no sentido de se pedir uma autorização para colocar esses conteúdos em acesso livre.
- b) Se a biblioteca não conseguir obter autorização para a disponibilização total de uma determinada revista, pode-se chegar a um acordo com os seus editores e autores para disponibilizar apenas parte da revista servindo os interesses de ambos.

No caso da revista *Philosophica*, por exemplo, esta premissa já foi colocada. A biblioteca possui um prazo mínimo de cinco anos de *delay* para aquela revista, ou seja, serão digitalizadas todos os exemplares, compromete-se a biblioteca a não colocar na Internet os últimos cinco números. Por outras palavras, será disponibilizado apenas o histórico da revista. Esta foi uma exigência dos responsáveis da revista e que a biblioteca obviamente respeitou.

- c) Outra possibilidade definida pela biblioteca prende-se com a colocação de uma nota adstrita às imagens aquando do seu aparecimento no sítio da biblioteca digital expressando: “*se é autor desta obra e não pretende que o seu trabalho esteja em livre acesso, contacte-nos.*”
- d) Poderá igualmente ser necessário restringir o acesso apenas a estudantes registados na biblioteca, investigadores ou mecenas da instituição, se a revista assim o entender.

Se necessária, a restrição de acesso às obras é passível de ser realizada. A biblioteca digital da FLUL tem a capacidade de o fazer através da plataforma da Universidade de Lisboa que acolhe as imagens digitalizadas. No *interface* da biblioteca não existe forma de gerir estas colecções mas podem ser exportadas para o digitUL. Desta forma, podemos colocar em acesso livre, no *interface* da biblioteca, as revistas que o permitirem, e aquelas que não concederam a autorização para serem disponibilizadas, mantê-las-emos em acesso restrito na plataforma.

3.2.6.1 – PROCESSO DE PRÉ-DIGITALIZAÇÃO

Antes de se iniciar o processo de digitalização é necessário prepará-lo, realizando uma análise cuidada da colecção que pretendemos digitalizar e posteriormente disponibilizar via biblioteca digital. A esta análise denominámos de “processo de pré-digitalização” pois é este processo que levará à selecção e efectiva constituição dos exemplares que serão digitalizados.

3.2.6.1.1 - CRITÉRIOS DE SELECÇÃO DOS EXEMPLARES

Como já foi referido no ponto 3.2.2 deste capítulo, a instituição optou por disponibilizar em linha esta colecção específica de periódicos devido às razões já evocadas, no entanto, resolvida a questão da selecção das colecções a digitalizar, há que resolver outra questão relativa à selecção dos próprios exemplares dessa colecção que serão sujeitos à digitalização.

Uma colecção de publicações periódicas é composta por números, e devemos verificar se todos estão disponíveis e, se estiverem, qual deles é passível de ser guilhotinado para ser posteriormente digitalizado. Para concretizar esta tarefa foram definidos critérios de selecção relacionados com a escolha das obras a digitalizar dentro daquele universo de revistas já seleccionadas. Os critérios de selecção para a digitalização foram os seguintes:

- a) Escolher o exemplar que estiver em melhor condição para digitalizar.
- b) Guilhotinar⁷ apenas exemplares que existam em duplicado e preferencialmente em triplicado na biblioteca;

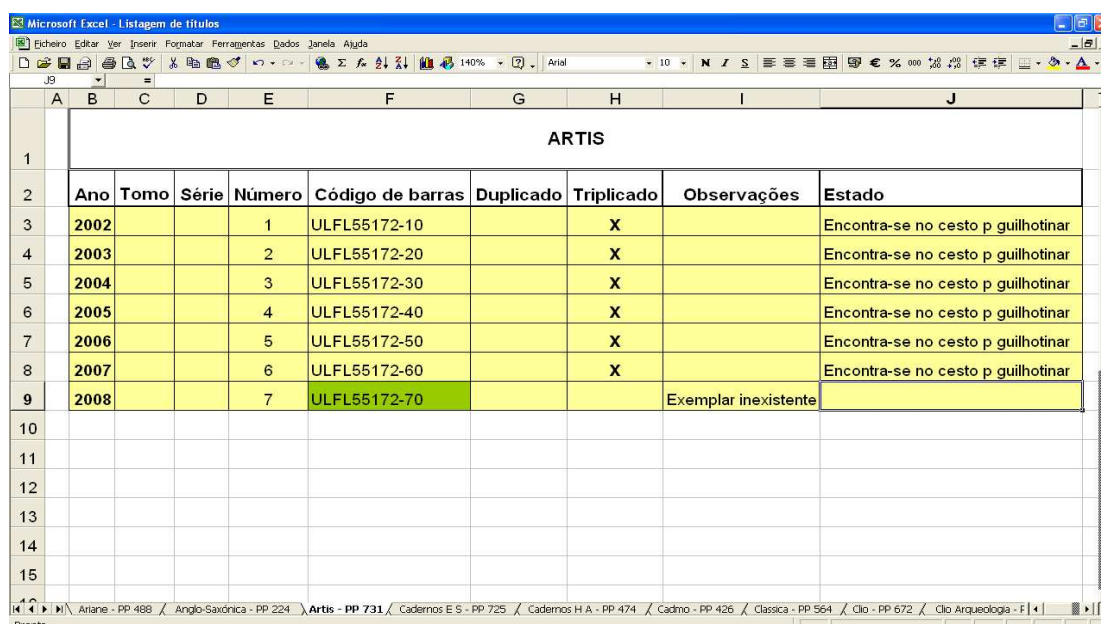
A biblioteca definiu estes dois critérios de selecção para colmatar duas situações: o primeiro pretende solucionar o problema da falta de exemplares, estabelecendo primazia à “guilhotinagem” de números que existam em duplicado ou em triplicado, e só em último caso é que se guilhotinarão exemplares únicos, que posteriormente serão reencadernados. O segundo critério, pretende responder à opção da biblioteca pela digitalização dos exemplares que se encontrem nas melhores condições físicas, no sentido de se obter imagens com qualidade e elevada definição que retratem o original a ser posteriormente disponibilizado via Internet.

⁷ Os exemplares seleccionados sofriam um corte preciso na lombada de modo a permitir a sua digitalização através do *scanner* com alimentador.

3.2.6.1.2 – O PROCESSO DE LEVANTAMENTO

No sentido de verificar as verdadeiras condições da colecção que pretendemos digitalizar, foi realizado um levantamento nos depósitos da biblioteca que consistiu no seguinte: As revistas foram pesquisadas no catálogo colectivo SIBUL (Sistema Integrado das Bibliotecas da Universidade de Lisboa) onde se verificou o número de existências e o seu estado. De seguida, confirmou-se se a informação disponível no SIBUL coincidia com as existências e o estado dos exemplares que se encontravam nos depósitos da biblioteca.

Após este levantamento, foram verificados os números que existiam em duplicado e triplicado. Registámos todas as informações recolhidas durante este processo num ficheiro Excel, criado para este propósito, denominado “*Listagem de Títulos*”, onde preenchimos os campos: *Ano*, *Tomo*, *Série*, *Número*, *Código de barras*, *Duplicado*, *Triplicado*, *Observações* e *Estado* para cada revista. Este ficheiro resumia o estado em que se encontravam as revistas que seleccionámos para digitalizar, como por exemplo: se seria necessário adquirir mais exemplares, se havia algum em falta, se dispúnhamos de exemplares em triplicado como desejável, etc.



The screenshot shows a Microsoft Excel spreadsheet with the title bar 'Microsoft Excel - Listagem de títulos'. The spreadsheet has columns labeled A through J. Column A contains row numbers 1 through 15. Column B is labeled 'Ano', C is 'Tomo', D is 'Série', E is 'Número', F is 'Código de barras', G is 'Duplicado', H is 'Triplicado', I is 'Observações', and J is 'Estado'. The data is for the journal 'Artis'. Rows 3 through 9 contain data for the years 2002 to 2008. The 'Duplicado' column is empty for all rows. The 'Triplicado' column contains 'X' for rows 3 through 8 and is empty for row 9. The 'Observações' column contains 'Encontra-se no cesto p guilhotinar' for rows 3 through 8 and 'Exemplar inexistente' for row 9. The 'Estado' column is empty for all rows.

	A	B	C	D	E	F	G	H	I	J
1										
2		Ano	Tomo	Série	Número	Código de barras	Duplicado	Triplicado	Observações	Estado
3		2002			1	ULFL55172-10		X		Encontra-se no cesto p guilhotinar
4		2003			2	ULFL55172-20		X		Encontra-se no cesto p guilhotinar
5		2004			3	ULFL55172-30		X		Encontra-se no cesto p guilhotinar
6		2005			4	ULFL55172-40		X		Encontra-se no cesto p guilhotinar
7		2006			5	ULFL55172-50		X		Encontra-se no cesto p guilhotinar
8		2007			6	ULFL55172-60		X		Encontra-se no cesto p guilhotinar
9		2008			7	ULFL55172-70			Exemplar inexistente	
10										
11										
12										
13										
14										
15										

Figura 1 – Ficheiro “Listagem de Títulos” – Revista *Artis*

Como podemos verificar na figura 1, a revista *Artis* possuía todos os exemplares em triplicado, à excepção do último que tem a nota de “*exemplar inexistente*”, isto porque é um exemplar que ainda não deu entrada na biblioteca. No entanto, esta revista foi uma excepção. A maior parte da colecção possuía muitos números marcados como “exemplares únicos” e alguns inexistentes na biblioteca. Estas lacunas teriam de ser colmatadas, antes de se iniciar a digitalização.

Desta forma, a biblioteca definiu as seguintes directrizes:

- a) Existindo exemplares em triplicado, colocamos essa informação no ficheiro e seleccionamos um deles para ser posteriormente guilhotinado. Como mencionámos, serão escolhidos aqueles que apresentarem as melhores condições de preservação;
- b) No caso de existirem exemplares em duplicado ou apenas um exemplar, devemos colocar essa informação no ficheiro, e sinalizar a vermelho esses números;
- c) Relativamente aos números duplicados, deve verificar-se no depósito da Secção das Obras Duplicadas se existem mais exemplares. Se existirem, a revista passa a triplicado, sendo esses números passíveis de “guilhotinagem”. Essa informação deverá constar no ficheiro de Excel. Se não existirem mais exemplares, levamos o exemplar que reunir melhores condições para ser sujeito ao processo de “guilhotinagem”, ficando apenas um exemplar intacto.
- d) Quando possuímos apenas um exemplar, devemos colocar uma nota de alerta para que se procure outro exemplar do mesmo número. Contudo, se isto não for possível, será necessário guilhotiná-lo e posteriormente reencaderná-lo.
- e) A confirmação do número das existências deverá ser realizada também na sala de leitura, ou seja, temos de verificar se os últimos exemplares das revistas estão presentes na sala de leitura, tendo em conta que, as revistas de tiragem anual devem ter disponíveis os últimos 5 números e as semestrais, os últimos 10 números. Essa informação também deverá constar no ficheiro.

Definida esta situação, os títulos que respeitavam todas as condições estipuladas pela instituição seguiram para a fase da “guilhotinagem”, sendo colocados numa caixa identificada para o efeito e posteriormente guilhotinados nas reprografias existentes na Faculdade. Para visualizar todo o processo de levantamento aqui exposto, ver o documento em anexo: ANEXO 3 – Processo de levantamento.

Após concluirmos este levantamento, decidimos começar o projecto com a digitalização da revista *Artis – Revista do Instituto de História da Arte da Faculdade de Letras de Lisboa*. Depois de se realizar um levantamento acerca do estado destas colecções, concluiu-se que esta era a única revista totalmente completa, existindo na biblioteca todos os números e em triplicado, o que possibilitava a “guilhotinagem” de um exemplar de cada número sem qualquer constrangimento.

A revista *Artis* começou a ser editada em 2002, aquando da publicação do seu primeiro volume. Trata-se de uma revista de periodicidade anual que possui, actualmente, uma colecção composta por 6 números.

3.2.6.2 – O PROCESSO DE DIGITALIZAÇÃO

Concluída a análise do espólio da biblioteca, a selecção da colecção e a definição dos documentos a digitalizar, transitamos para o processo de digitalização, propriamente dito.

Uma vez que a colecção a digitalizar era constituída por publicações em série, ou revistas, era natural que as imagens provenientes contivessem maioritariamente mancha de texto mas também muitas imagens e ambas teriam de transcrever a qualidade máxima do original. Com este propósito, seguimos as orientações definidas no ponto 2.4.3, do Cap. II relativas à digitalização e aos requisitos de qualidade das imagens.

3.2.6.2.1 – RECURSOS TÉCNICOS

Como referimos no ponto 2.4.2, do Cap. II, a implementação de um processo de digitalização depende de recursos técnicos que possibilitam a sua concretização.

Neste projecto, os equipamentos utilizados foram:

- Dois *hardware* CPU com um processador *Intel(R) Pentium(R) Dual CPU (E2220 @2.40GHz. 2.40GHz, 1,99 GB de RAM)* e sistema operativo *Microsoft Windows XP. Profissional, Versão 2002*;
- Dois monitores com ecrã LCD de 24 polegadas;
- Um disco externo;
- Dois scanners da marca *Scanner Kodak i1220* (scanner com alimentador e scanner de mesa);
- Uma “cunha de cor”- *TIFFEN Color Separation Guide and Gray Scale*;
- Os *Softwares*: ScandAll, PAPAIA e ContentE.

O hardware era completamente compatível com os dispositivos de entrada e saída e softwares utilizados.

Os dois *scanners* eram adequados a este processo de digitalização em particular. O *scanner* com alimentador permitiu a digitalização das folhas soltas de toda a obra previamente guilhotinada, e o *scanner* de mesa, por sua vez, possibilitou a digitalização das capas e contra-capas das revistas. (ANEXO 2, Figura 1 e 2)

As capas foram digitalizadas juntamente com uma “cunha de cor”, fundamental para recuperar posteriormente as cores das imagens digitalizadas. A cunha de cor TIFFEN

Color Separation Guide and Gray Scale (ANEXO 2, Figura 3) é um instrumento de controlo de qualidade de tons com duas componentes: a escala de cinzentos e a de controlo de cor. Esta “régua” permite responder a possíveis problemas de preservação futura da imagem, uma vez que é através dela e por meio de comparação de cores que podemos chegar ao tom original da imagem.

Durante o processo de digitalização das revistas foi necessário perceber quanto espaço de disco era necessário para “alojar” as imagens resultantes da digitalização. A pasta existente no servidor da biblioteca não comportava o espaço necessário para a cópia de ficheiros das imagens existentes e a criar, assim, foi necessário solicitar ao departamento de informática que se cria-se uma pasta com mais espaço em disco para abarcar o projecto.

O disco externo funcionava como *backup* de apoio e segurança de modo a assegurar que aquelas imagens não se perdiam por algum problema existente no servidor da biblioteca. Aqui foram gravadas as revistas completamente terminadas e prontas para a disponibilização.

3.2.6.2.2 - SOFTWARE UTILIZADO – SCANDALL 21

As obras foram digitalizadas através do programa *ScandAll 21*, programa adquirido aquando da compra do *Scanner Kodak i1220* e as imagens visualizadas através do programa de visualização de imagens *ACDSee Pro Photo Manager*.

3.2.6.2.3 - EXECUÇÃO

A execução do processo de digitalização foi facilitada em parte pela estrutura linear das obras que compunham este projecto. De facto, falamos de uma colecção de obras relativamente recentes (à excepção da *Revista da Faculdade de Letras* cujo primeiro número data de 1930) e cuja estrutura é basicamente linear, sem grandes alterações de formato, sem necessidades de conservação, o que possibilitou uma digitalização efectiva e o desenvolvimento *In-house* deste projecto. Esta estrutura permitiu que os exemplares seleccionados fossem guilhotinados para facilitar o processo de digitalização.

As obras foram digitalizadas em formato TIFF, com uma profundidade de cor de 24 bits e resolução de 300 dpi. As imagens daqui resultantes foram gravadas com a designação do código de barras da revista a que pertenciam e por número sequencial de imagem, exemplo: ULFL55172-10_0001. Seguimos, desta forma, as regras para a denominação inequívoca de ficheiros, como definido no ponto 2.4.3, no capítulo II.

A gravação das imagens era feita dentro de pastas criadas para o efeito e guardadas dentro de um domínio da Biblioteca digital da FLUL denominada de “Guardar Imagens”

com o nome da revista e respectivo código de barras, cujo caminho era C:\GUARDAR IMAGENS\Artis\ULFL55172-10 (exemplo da revista Artis, n.º 1).

A execução do processo de digitalização obedeceu a duas fases: a fase da digitalização das capas efectuadas no *scanner* de mesa, e a fase da digitalização do conteúdo da obra⁸, realizadas no *scanner* com alimentador. (ANEXO 4 – Processo de digitalização).

1. DIGITALIZAÇÃO DAS CAPAS E CONTRA-CAPAS

Para iniciar o processo de digitalização abríamos o programa *ScandAll 21* e seleccionámos o *scanner flatbed*. As capas e contra-capas foram introduzidas manualmente no scanner, posteriormente realizou-se um *preview* onde se verificou se as mesmas estavam perfeitamente alinhadas com as margens, assim sendo, foram digitalizadas e guardadas na respectiva pasta da revista a que pertenciam.

As capas, e apenas estas, foram digitalizadas com uma “cunha de cor” *TIFFEN Color Separation Guide and Gray Scale*. Contudo, será importante referir que por opção da biblioteca, esta “cunha de cor” não será apresentada posteriormente aquando da disponibilização das imagens *On-line*. Devido a esse facto, houve a necessidade de deixar uma margem de 1cm entre a régua e a capa, para possibilitar que a régua fosse “cortada” aquando da disponibilização da imagem.

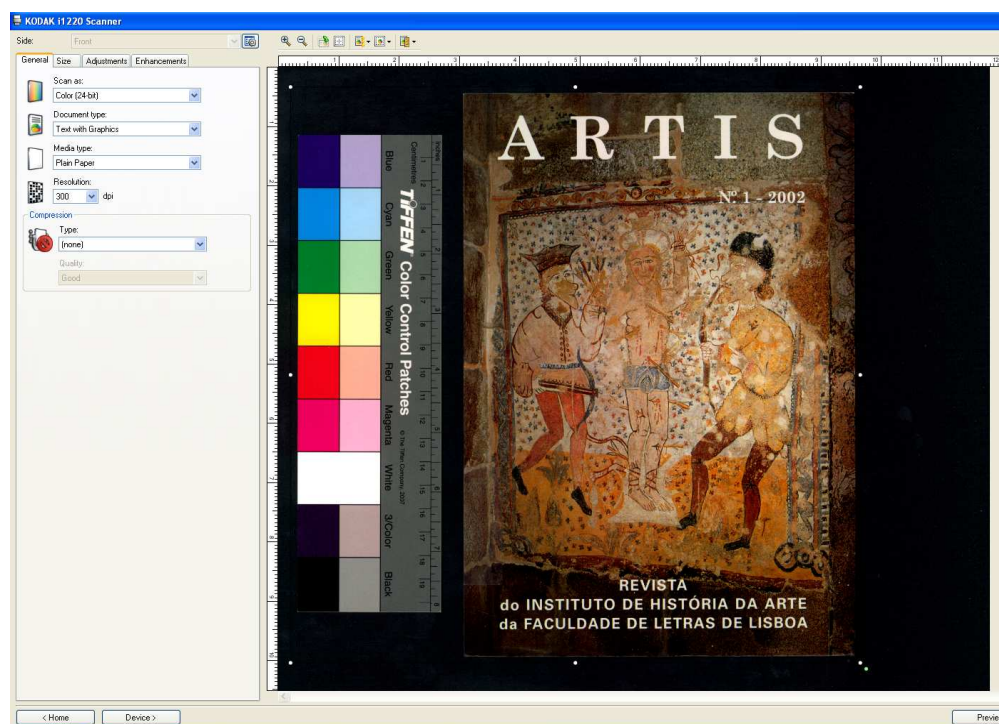


Figura 2 – Exemplo digitalização de capas e contra-capas

⁸ A parte interna da obra revela a fracção intelectual ou o conteúdo do livro. É através dele que a obra é considerada um objecto cultural.

2 - DIGITALIZAÇÃO DO CONTEÚDO DA OBRA

Para a digitalização do conteúdo da obra utilizámos o modelo *Bib digital ADF* correspondente ao *scanner* com alimentador automático. As páginas guilhotinadas da obra foram colocadas no alimentador seguindo a ordem de disposição na obra e após concluído o processo, foram guardadas na sua pasta correspondente.

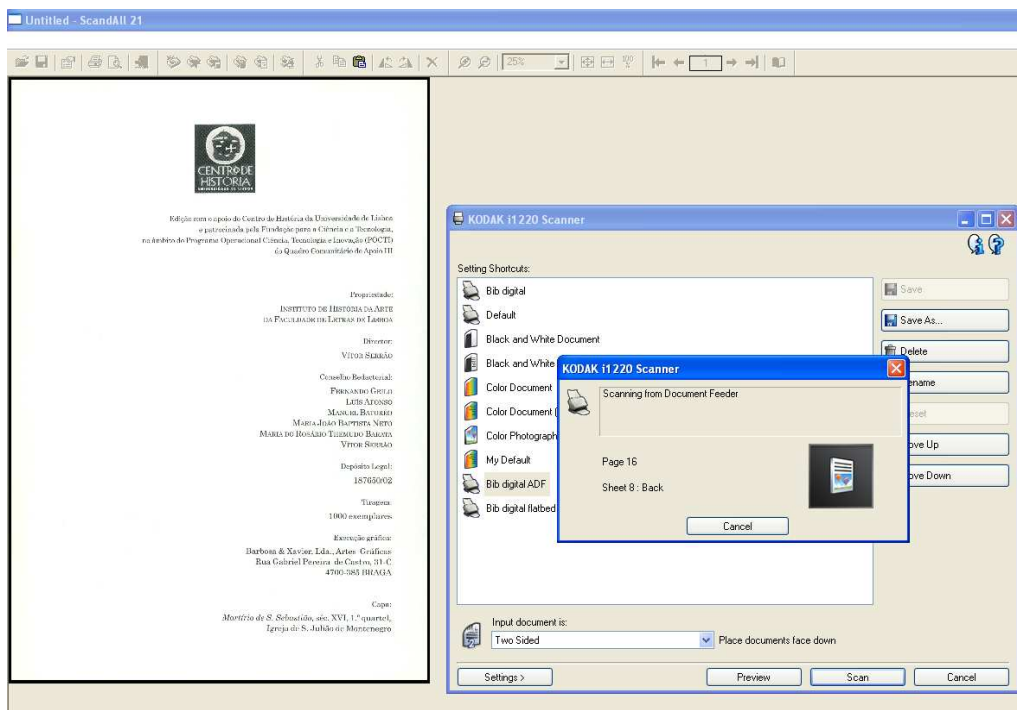


Figura 3 – Exemplo digitalização do conteúdo da obra

Todavia, já nesta fase, foi necessário fazer um controlo de qualidade, ainda que de modo superficial, sobre as imagens pois qualquer impureza que aparecesse na lente do *scanner* poderia originar pequenos “traços” ou manchas nas imagens.

3.2.6.3 – O PROCESSO DE IMPLEMENTAÇÃO DE METADADOS

Como já referimos no ponto 2.4.3. do Cap. II, quando desenvolvemos a temática dos metadados, estes “dados sobre dados” que introduzimos nos ficheiros digitais constituem, provavelmente, a facção mais importante num projecto de biblioteca digital, como tal, a sua introdução tem de ser rigorosa pois qualquer falha pode pôr em causa todo o projecto.

3.2.6.3.1 - SOFTWARE UTILIZADO – PAPAIA

A nomeação das imagens e a introdução dos metadados foi feita através do programa PAPAIA, uma aplicação desenvolvida pela BN que permite fazer o tratamento das obras na fase posterior à digitalização das mesmas.

As principais funcionalidades deste programa são a renomeação automática de ficheiros de imagem (cota da obra, número sequencial, data de edição da obra, etc.), a edição de metadados TIFF e a estruturação das imagens. O software permite ainda que adicionemos dados que caracterizam a obra digitalizada tanto ao nível técnico como descritivo. Para além da normalização e estruturação das imagens com este programa é possível criar um ficheiro XML (eXtensible Markup Language) que será utilizado posteriormente no processo de disponibilização para criação dos índices de navegação da obra.

3.2.6.3.2 - EXECUÇÃO

O primeiro passo será importar as imagens que queremos tratar para o referido programa. Antes de iniciarmos a nomeação das imagens, necessitamos de configurar os dados que pretendemos embutir nas mesmas através do preenchimento de vários campos existentes no *software*.

No campo *Images Header*, insere-se a informação relativa ao cabeçalho das imagens. Neste campo editámos os cabeçalhos TIFF segundo a tabela seguinte:

Designação	Etiquetas TIFF	Valores (Conteúdo)
1. ImageWidth	256	(N.º de pixels na horizontal)
2. ImageLength	257	(N.º de pixels na vertical)
3. BitsPerSample	258	8, 8, 8 (Cor)
4. Compression	259	1 (Sem compressão)
5. Photometric Interpretation	262	2 (Cor, default para RGB)
6. Fillorder	266	1 (Default)
7. Document name	269	ULFL.....-..
8. Make	271	Kodak
9. Model	272	i1220
10. StripOffsets	273	0 (byte offset de cada strip)
11. Orientation	274	1 (Orientação da imagem, normal)
12. Samplesperpixel	277	3 (Default para RGB)
13. RowsPerStrip	278	(N.º de linhas por strip)
14. StripByteCount	279	Não preencher, não há compressão
15. Xresolution	282	(N.º de pixels/unid. resol. na horizontal)
16. Yresolution	283	(N.º de pixels/unid. resol. na vertical)
17. Planarconfiguration	284	1
18. Resolution unit	296	2 (Polegadas)
19. Software	305	ScandAll 21
20. DateTime	306	(Data e hora de captura, formato YYY:MM.DD HH:MM:SS)
21. Artist	315	DBFLUL
22. HostComputer	316	Intel(R) Pentium(R) Dual CPU. E2220 @2.40GHz. 2.40GHz, 1,99 GB de RAM Microsoft Windows XP. Profissional. Versão 2002. Service Pack 3
23. Copyright	3432	ULFL, 2009

Tabela 2 – Etiquetas TIFF utilizadas

Esta tabela indica as etiquetas TIFF correspondentes a dados relativos à imagem digital. Neste projecto, colocámos manualmente a descrição das etiquetas que se encontram na tabela a cinzento, com a descrição seguinte:

Artist – DBFLUL

Copyright – ULFL, 2009

Doc name – “Código de barras da obra”

Fill order – 1

Host computer – “modelo de computador utilizado”

Make – “marca do scanner” - *Kodak*

Model – “modelo do scanner” - i1220

Software – ScandAll 21

Orientation – 1

Como podemos verificar, a descrição destas etiquetas permite identificar quem produziu as imagens e quando, qual é o código de barras da revista à qual aquela imagem pertence e qual o modelo de computador, *scanner* e *software* que foi utilizado para produzir as mesmas. (ANEXO 5, fig. 5)

Seguidamente aplicamos outros metadados, estes referentes ao campo *Image Numbering* onde se insere a informação relativa à numeração das imagens. (ANEXO 5, fig. 7) Neste projecto, alteramos os campos seguintes:

- Fórmula: %c_%d_%i_%p_t0 – Esta fórmula é a sintaxe desejada para a escrita dos nomes quando se trata de periódicos, sendo que %c corresponde à cota, %d à data, %i ao número das imagens e %p ao número da página a que corresponde no original.

- *Book properties*:

Prefixo: código de barras da revista;

Images directory: o caminho da directoria onde se encontram as imagens;

Data: data da publicação;

Custom: permite adicionar novos valores ao tipo de numeração, neste projecto, adicionámos a designação “b” correspondendo à descrição de página branca.

Depois de aplicarmos estes dados, procedemos à renomeação das imagens e à sua consequente estruturação. Utilizámos uma sintaxe definida pela BND para nomear imagens resultantes de processos de digitalização de originais. A sintaxe apropriada para as publicações periódicas, alvo do nosso trabalho, consistia no seguinte:

<cota_da_obra> _ <data> _ <ordem_da_imagem> _ <referência_da_página> _ t0
 ex: ULFL55172-10_2002-10_0003_rosto_t0.tif

Como podemos verificar, são atribuídos às imagens os seguintes elementos base:

- O item <cota_da_obra> é um identificador único que se deve manter constante para todas as imagens da mesma obra. Optámos por identificar as nossas imagens não com o identificador “cota da obra” mas sim com o “código de barras” da mesma;
- O elemento <data> um factor relevante no caso de publicações periódicas, optámos por referir o ano e o mês da publicação;
- O item <ordem_da_imagem> representa o número de sequência original da imagem digitalizada;
- O item <referência_da_página> atribui à imagem o número correspondente à paginação que tem na obra. No caso das imagens sem numeração, podemos nomeá-las como “capa”, imagem “branca”, “rosto”, entre outros;
- O valor <t0.tif> é um elemento que distingue imagens originais, as chamadas “masters” de outras versões, corresponde portanto às imagens em formato TIFF.

A figura seguinte é exemplificativa deste processo:

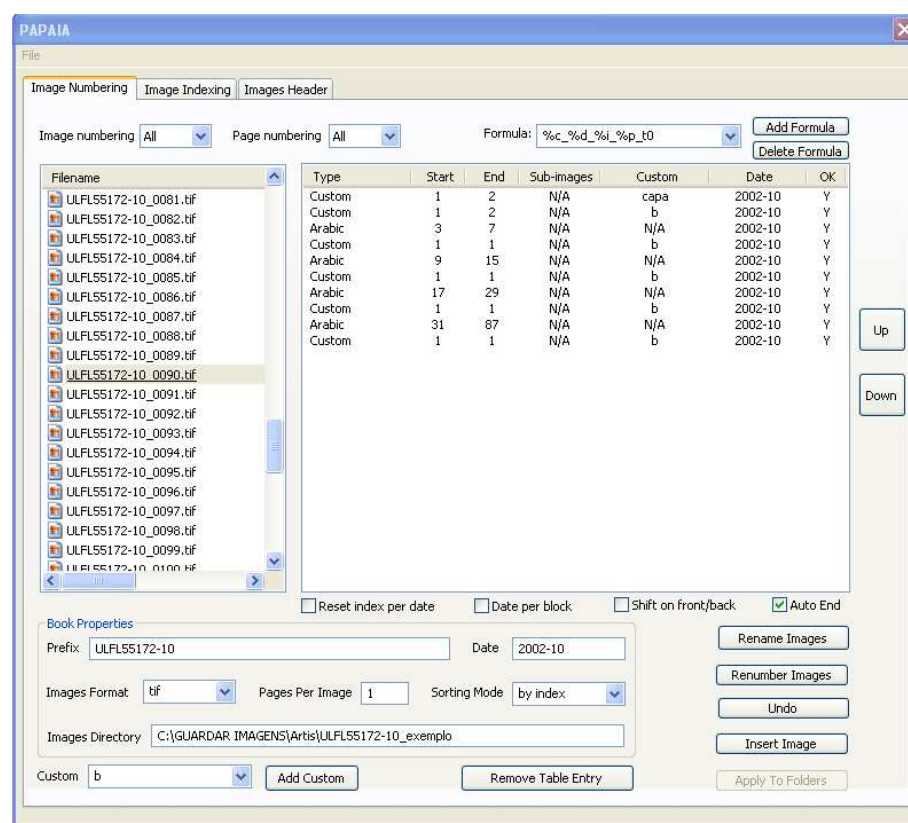


Figura 4 – Exemplo da normalização e estruturação das imagens – PAPAIA

Terminado o processo de introdução de metadados e nomeação de ficheiros para todas as imagens da obra, passamos para a fase seguinte que corresponde à criação de um ficheiro em formato XML gerado automaticamente pelo PAPAIA. Este documento XML serve para estruturar, armazenar e trocar informação e será posteriormente utilizado pelo programa ContentE no processo de disponibilização. (Para analisar todo o processo de implementação de metadados e nomeação de imagens, ver ANEXO 5)

3.2.6.4 - CONTROLO DE QUALIDADE DAS IMAGENS DIGITALIZADAS

Antes de disponibilizar as revistas já concluídas (digitalizadas e embutidas de metadados), a biblioteca decidiu fazer um controlo de qualidade rigoroso no sentido de ter a certeza de que não existia nenhuma lacuna e que todas as imagens transcreviam fielmente o documento original. (ANEXO 6 – Controlo de qualidade).

O controlo de qualidade foi realizado de uma forma escalonada segundo a premissa de que a pessoa que tinha digitalizado uma revista não podia fazer o controlo de qualidade da mesma e que haveria sempre um controlo final efectuado pela principal responsável pelo projecto que daria a revista como concluída. Assim sendo, uma revista só seria dada como definitiva/concluída depois de passar pelo controlo de qualidade realizado por três pessoas.

Os campos verificados neste controlo de qualidade foram os seguintes: identificação da revista, qualidade da imagem digital, qualidade da nomeação das imagens, qualidade dos metadados e gravação da revista no disco externo.

Através deste método foram encontradas algumas anomalias que tinham passado despercebidas anteriormente no processo de digitalização e implementação de metadados e que obviamente teriam de ser corrigidas. (ANEXO 6, fig. 3) Se o erro se prendia unicamente com a nomeação das imagens, teríamos de refazer a revista toda no programa PAPAIA, todavia, se o erro se prendia com uma anomalia na digitalização, teríamos de digitalizar de novo essas imagens e ainda refazer os metadados da totalidade da revista no programa PAPAIA.

Findo o processo de controlo de qualidade e depois de todas as correcções necessárias, seria guardada num disco externo por uma questão de segurança uma cópia das revistas já totalmente correctas.

3.2.6.5 – O PROCESSO DE DISPONIBILIZAÇÃO

3.2.6.5.1 - O SOFTWARE UTILIZADO – ContentE

A aplicação utilizada para estruturar os ficheiros, de modo a disponibilizá-los em linha, foi o software livre ContentE (editor de conteúdos estruturados) desenvolvido pela Biblioteca Nacional. Este programa possibilita a construção de estruturas de obras digitalizadas e através dele podemos criar uma multiplicidade de índices de visualização de uma obra segundo as necessidades de cada instituição.

3.2.6.5.2 – EXECUÇÃO

A realização do processo de disponibilização constituiu várias etapas. Primeiro criámos uma “obra” importando as imagens TIFF já estruturadas e nomeadas pelo PAPAIA, para este programa. De seguida, anexámos o registo bibliográfico da revista que estávamos a tratar a esta obra, tendo o cuidado prévio de converter o registo para um formato que o ContentE lê, como o TXT⁹ ou um registo XML.

De facto, a disponibilização de imagens através este programa é feita por imagens JPEG, que como já referimos no ponto 2.4.3 do Cap. II, é um formato que alia uma boa qualidade de imagem a um menor peso em disco. Desta forma, é necessário converter as imagens TIFF em JPEG ou PDF, este *software* permite fazer automaticamente essa conversão.

Sendo que as imagens TIFFs constituem ficheiros de arquivo, é necessário fazermos a conversão para estes dois formatos porque serão eles os utilizados na disponibilização ao público.

Os ficheiros JPEG foram criados com resolução de 150dpi e 90 de qualidade (numa escala de 0-100) seguindo as boas práticas da BN neste campo. Foram criadas também miniaturas neste formato mas apenas das páginas de rosto, com largura, de 140 *pixels*. Estas miniaturas serão apresentadas no sítio da biblioteca digital como exemplo da revista a visualizar.

A biblioteca optou colocar disponível *On-line* apenas as imagens em formato JPEG, as imagens em formato PDF ficarão em arquivo e apenas serão disponibilizadas a uma pessoa particular ou a uma instituição se for feita uma solicitação com as razões do pedido bem fundamentadas. Estas restrições pretendem salvaguardar os direitos de propriedade do material disponibilizado.

⁹ A extensão .txt é um formato de texto plano, simples, sem recursos a formatação que pode ser lido ou aberto por qualquer programa que lê texto.

Após a criação dos ficheiros JPEG é possível criar os índices que aparecerão posteriormente na página da biblioteca digital. A biblioteca optou por utilizar o mesmo tipo de índice usado no projecto anterior e que podemos visualizar na figura seguinte:

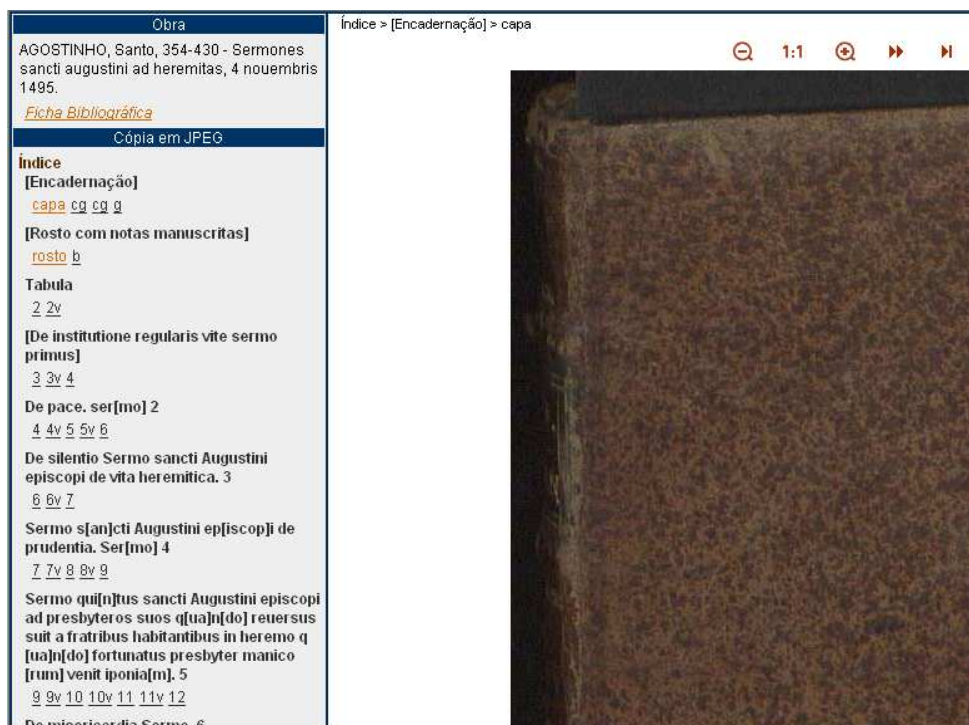


Figura 5 – Tipo de índice utilizado na biblioteca digital da FLUL através do ContentE

No sentido de concebermos um novo índice para a obra que queremos tratar, é necessário criar uma estrutura com capítulos no índice através de um sistema de nós e posteriormente através de um sistema “*drag and drop*” introduzir as imagens JPEG no respectivo capítulo. (ver fig. 10 do Anexo 7)

Para além de vários tipos de índices, o programa ContentE disponibiliza uma lista de estilos pré-definidos que contêm as definições de estilo de várias instituições com a especificação da cor que identifica a instituição, os seus direitos de *copyright*, entre outros. O mesmo se passa para a biblioteca da FLUL que dispõe de um índice já pré-definido.

Os exemplos seguintes revelam essa personalização do *layout* de disponibilização através dos estilos. Na versão da BND, o *layout* de disponibilização é vermelho e a menção que apresenta referente aos direitos de *copyright* é a seguinte: “© Biblioteca Nacional Digital – ContentE v.1.6”:

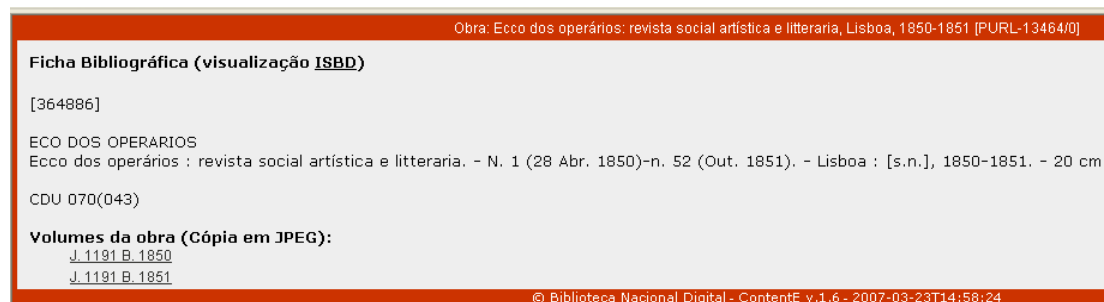


Figura 6 – Layout da BND

No que concerne ao estilo do *layout* de disponibilização da biblioteca da FLUL, o seu fundo é azul (o azul específico da FLUL), com a designação referente aos direitos de *copyright* “© Faculdade de Letras da Universidade de Lisboa – ContentE v.1.6”:

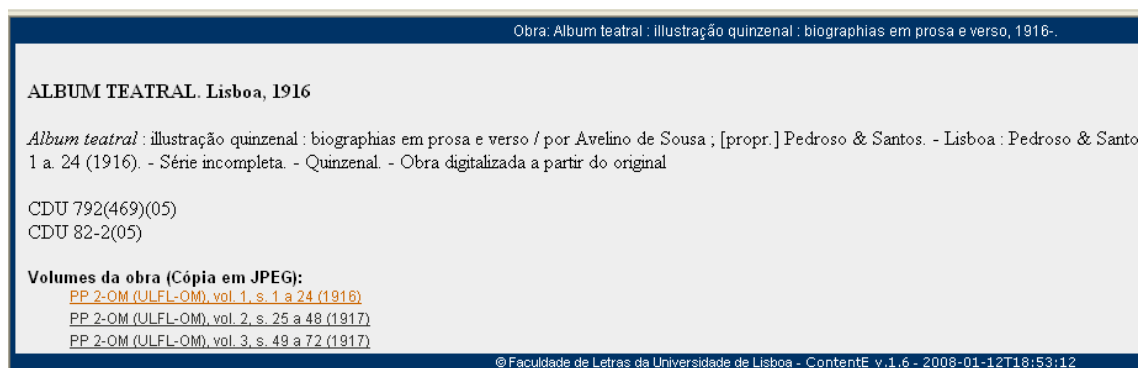


Figura 7 – Layout da Biblioteca Digital da FLUL

Na verdade, um estilo pré-determinado não é definido apenas pela cor da instituição ou designação de direitos de autor. Ele atribui à futura página de disponibilização das imagens outras características, como o tipo de “botões” apresentados e as *style sheets* (folhas de estilo) apropriadas para cada instituição. As “folhas de estilo” contêm as instruções que controlam a apresentação de documentos num monitor, numa impressora ou outros dispositivos, dando-lhes uma estrutura. Assim, todas estas características presentes nos estilos predefinidos do ContentE personalizam uma determinada página *Web* de apresentação de conteúdos digitais.

De facto, o processo de disponibilização que aqui apresentamos é apenas uma demonstração de como pode ser utilizado, em traços gerais, o programa ContentE, isto porque, e ao contrário dos outros processos apresentados ao longo deste trabalho, não foi concluído a tempo da finalização deste trabalho-projecto. (Para visualizar o processo de disponibilização através do programa ContentE, ver o ANEXO 7).

3.3 - RESULTADOS

Será importante reflectir sobre os efectivos resultados deste projecto. Na verdade, a meta que inicialmente fora estipulada, e que consistia na digitalização de seis títulos até ao final do ano de 2009, não foi atingida. Até Agosto de 2009, apenas tinha sido digitalizada uma revista na sua totalidade, a revista *Artis*, e grande parte da revista *Philosophica*.

Isto deveu-se à conjugação de diversos factores. Por um lado, sucederam-se problemas com o funcionamento dos *scanners*, devido ao aparecimento de manchas nas imagens digitalizadas. Tal circunstância causou a repetição do processo de digitalização em alguns exemplares. Por outro lado, os recursos humanos incorporados no projecto, uma estagiária em regime pós-laboral e um número de alunos voluntários, não permitiram um desenvolvimento célere devido ao pouco tempo a ele atribuído. A presença de voluntários levou ainda a que se consumisse grande parte do tempo em consecutivas acções de formação, o que originava também atrasos.

O incumprimento da meta traçada deveu-se também à abordagem de planeamento do projecto. A meta fora calculada tendo em conta a presença constante de voluntários, que contribuíam com um determinado número de horas pré-acordado. Tal não foi cumprido.

Um outro factor foi também preponderante para o incumprimento verificado. Falamos das “falhas” registadas durante a digitalização de alguns exemplares e durante a implementação dos metadados. A maior parte dos recursos humanos nunca tinha participado num projecto desta natureza. Assim, registavam-se frequentemente erros que tinham de ser solucionados. Em consequência disto, a revista *Artis* viu os seus metadados serem totalmente refeitos; e o mesmo aconteceu com alguns exemplares da revista *Philosophica* no que respeita à digitalização.

Tão pouco foi possível disponibilizar a revista concluída via biblioteca digital, devido a um problema técnico que se prende com a estruturação de alguns parâmetros existentes no programa ContentE.

Como tínhamos previsto na análise SWOT, a ausência de *know-how* por parte dos técnicos da instituição para trabalhar com este programa teve consequências. A biblioteca viu-se obrigada a esperar por ajuda externa para resolver o problema de actualização do *software*, e aguarda neste momento a personalização, ou seja, a adaptação do programa ContentE às suas necessidades específicas.

O atraso verificado deveu-se também às características das “folhas de estilo” existentes no programa ContentE. As *style sheets* relativas à biblioteca digital da FLUL são da responsabilidade da empresa DigiCult, que desenvolveu a primeira fase do projecto de biblioteca digital, e não da dos técnicos da biblioteca. Assim, terá de ser feita uma

adaptação destas “folhas de estilo”, ou a criação de outras, que levem em conta as alterações necessárias.

A criação de novas *style sheets* só pode ser conseguida com a colaboração dos responsáveis do programa ContentE. Foi requerida ajuda neste sentido e esperamos uma resolução do problema com a maior brevidade possível.

Solucionados os problemas acima referidos, podemos avançar para a disponibilização de um periódico completo, a revista *Artis* e brevemente de outro periódico já praticamente em fase de conclusão, a revista *Philosophica*.

CAP. IV – CONCLUSÕES

Terminada a apresentação do projecto *Biblioteca digital da FLUL – 2ª fase. O Projecto de disponibilização de publicações periódicas* a que nos propusemos no início do presente trabalho, resta-nos fazer uma reflexão global e retirar algumas considerações gerais.

Este trabalho insere-se numa realidade em crescimento, o desenvolvimento de novos serviços de disponibilização de informação centrados no “novo” utilizador, como é o caso dos serviços de bibliotecas em linha. Os projectos de biblioteca digital, desenvolvidos em grande parte por bibliotecas universitárias, traduzem-se em novos serviços que têm, entre as suas valências, a de alcançarem mais facilmente utilizadores reais e potenciais.

O estudo que aqui desenvolvemos revela como foi implementado um projecto de biblioteca digital *In-house*, com poucos recursos humanos e financeiros, baseado fundamentalmente no sistema técnico existente na instituição e utilizando programas não-proprietários desenvolvidos pela BN. O estado de conservação das revistas, sem grandes alterações de formato e o facto de serem passíveis de “guilhotinagem” possibilitou a redução dos custos do processo de digitalização e o desenvolvimento deste projecto nas instalações da biblioteca.

A metodologia utilizada para a elaboração deste trabalho foi uma “metodologia projecto” definida em seis etapas. Começámos por definir o público-alvo deste projecto, seleccionámos a colecção a disponibilizar *On-line*, identificámos o problema que o projecto vinha resolver, realizámos um diagnóstico através de uma análise SWOT, definimos os objectivos e concluímos com a implementação do projecto propriamente dito.

Ao longo do trabalho, identificámos a “questão” que este projecto pretendia solucionar através das seguintes perguntas: “*Como implementar um projecto desta natureza, com estas características, numa biblioteca digital, tendo como antecessor um projecto já com provas dadas de qualidade e sucesso?*”, “*Como digitalizar as imagens, organizá-las de forma coerente e dotá-las de metadados, de forma a estruturá-las, permitindo a sua reutilização em qualquer contexto agora e no futuro?* E ainda, “*Como disponibilizar este conteúdo On-line?*”. No desenvolvimento do trabalho, respondemos a estas questões por intermédio da apresentação de todos os processos que, conjugados, permitiram a execução do projecto. Foram descritas e analisadas as fases de desenvolvimento do projecto que exemplificavam como digitalizar as publicações periódicas, como organizar as imagens através de uma denominação inequívoca e em

pastas apropriadas, como introduzir os metadados nas imagens através do programa PAPAIA e ainda como utilizar o *software* ContentE para a sua disponibilização *On-line*.

Os resultados revestem-se de dois aspectos, um negativo, o outro positivo. O aspecto negativo relaciona-se com o incumprimento da meta estipulada pela instituição, que consistia em digitalizar seis revistas até ao final do ano 2009. Como tivemos oportunidade de referir, o incumprimento desta meta deveu-se a uma conjugação de factores desfavoráveis: um planeamento adoptado pouco eficaz, a inexistência de recursos humanos em permanência no projecto, problemas na implementação do processo de digitalização e de metadados devido, em grande medida, à inexperiência dos recursos humanos (voluntários), e, por fim, a desadequação do programa ContentE aos requisitos actuais da biblioteca.

No que respeita à gestão dos recursos humanos envolvidos neste projecto, a experiência demonstrou a eventual necessidade da existência de uma estratégia para a coordenação do trabalho voluntário de forma a estabelecer uma maior responsabilização das pessoas envolvidas.

O aspecto positivo resultante deste trabalho foi a digitalização total de uma revista (*Artis*) por meio de um trabalho criterioso e rigoroso. Outro resultado positivo prende-se com a experiência adquirida. Como assinalámos na análise SWOT, a experiência do desenvolvimento de um projecto *In-house*, capacitou a biblioteca de valências que anteriormente não tinha, resultando em consolidação de conhecimentos que conferem mais autonomia à instituição. Este factor constitui uma oportunidade fundamental para a biblioteca se destacar neste domínio da oferta de conteúdos digitais.

Ficou patente que é possível desenvolver um projecto de biblioteca digital com estas características (*In-house*, com escassos recursos financeiros e humanos, apenas com os recursos técnicos da instituição), desde que exista rigor na execução e respeito pelas normas nacionais e internacionais definidas para projectos desta natureza. Assim sendo, tendo em conta os recursos que dispúnhamos, não será excessivo referir que os resultados foram positivos.

Um dos aspectos que, embora tendo sido considerado, não foi objecto de um desenvolvimento cuidado neste projecto foi a questão da preservação digital. Esta questão é premente nos dias de hoje. Se por um lado, as características únicas do formato digital permitem rápida e facilmente criar informação, disponibiliza-la e actualiza-la diariamente, por outro lado, verifica-se ainda a existência de algumas limitações no que diz respeito à sua preservação para que posteriormente possa ser recuperada perfeitamente sem perda de conteúdo.

Tendo em conta esta problemática, adoptámos pela utilização de programas não-proprietários, o que permitirá a alteração e adequação dos programas segundo as necessidades futuras, e dotámos as nossas imagens digitais com os metadados requeridos no sentido de salvaguardar a sua preservação a longo prazo.

Concluindo, apesar da meta inicialmente estipulada não ter sido cumprida e de aguardarmos ainda a referida alteração no software para disponibilizar parte da colecção, consideramos que o desenvolvimento deste projecto foi positivo. Apesar dos limitados meios técnicos, financeiros e humanos, o trabalho realizado é válido e respeita em rigor os princípios de qualidade e normas predefinidas.

CAP. V – BIBLIOGRAFIA

ANDREWS, Judith; LAW, Derek - *Digital libraries: policy, planning and practice*. Aldershot: Ashgate Publishing, 2004. 272 p.

ARMS, William Y – *Digital Libraries* [Em linha]. Massachusetts: MIT Press, 2001. [Consult. 05 Abr. 2009]. Disponível na [www:<URL: http://books.google.pt/books?id=pzmt3pcBuGYC&printsec=frontcover&source=gbs_summary_r&cad=0>](http://books.google.pt/books?id=pzmt3pcBuGYC&printsec=frontcover&source=gbs_summary_r&cad=0)

Biblioteca da FLUL: apresentação [Em linha]. [Consult. 24 de Fev. 2009]. Disponível na [www:<URL: http://www.fl.ul.pt/biblioteca/apresentacao.htm>](http://www.fl.ul.pt/biblioteca/apresentacao.htm)

Biblioteca Digital FLUL: História [Em linha]. [Consult. 24 de Fev. 2009]. Disponível na [www:<URL: http://www.fl.ul.pt/biblioteca/biblioteca_digital/index.htm>](http://www.fl.ul.pt/biblioteca/biblioteca_digital/index.htm)

Biblioteca Nacional Digital – *Aplicação PAPAIA: Processamento de páginas digitalizadas* [Em linha]. Fev. 2005. [Consult. 09 de Jul. 2008]. Disponível na [www:<URL: http://bnd.bn.pt/tec/papaia/papaia-manual.html#1>](http://bnd.bn.pt/tec/papaia/papaia-manual.html#1) (actualmente indisponível)

BORBINHA, José Luís Brinquete – *Bibliotecas digitais: o futuro através da biblioteca tradicional*. Lisboa: Instituto Superior Técnico, 2000. Tese de doutoramento.

BORBINHA, José Luís [et al.] – *A gestão de obras digitalizadas na BND* [Em linha]. [Consult. 25 Fev. 2009]. Disponível na [www:<URL: http://xata.fe.up.pt/xata2005/papersfinal/43.pdf>](http://xata.fe.up.pt/xata2005/papersfinal/43.pdf)

CALIMERA GUIDELINES: digitalization [Em linha]. [Consult. 15 Ago. 2008]. Disponível na [www:<URL: http://www.calimera.org/Lists/Guidelines%20PDF/Digitisation.pdf>](http://www.calimera.org/Lists/Guidelines%20PDF/Digitisation.pdf)

CARIGNAN, Yvonne [et al.] – *Best practice guidelines for digital collections at University of Maryland Libraries* [Em linha]. Maryland: Ed by Susan Schreibman, 2007. [Consult. 27 Set. 2008]. Disponível na [www:<URL: http://www.lib.umd.edu/dcr/publications/best_practice.pdf>](http://www.lib.umd.edu/dcr/publications/best_practice.pdf)

CHAN, Lois; LEU ZENG, Maria - *Metadata Interoperability and Standardization – A Study of Methodology, Part I* [Em linha]. 2006. [Consult. 5 Maio 2009]. Disponível na [www:<URL:http://www.dlib.org/dlib/june06/chan/06chan.html>](http://www.dlib.org/dlib/june06/chan/06chan.html)

DEEGAN, Marilyn; TANNER, Simon – *Digital Preservation*. London: Facet Publishing, 2006. 260 p. (Digital futures series).

Digital Libraries: Future Directions for a European Research Programme [Em linha]. San Cassiano, 2001. [Consult. 15 Dez. 2008]. Disponível na [www:<URL:http://delos-noe.iei.pi.cnr.it/activities/researchforum/Brainstorming/brainstorming-report.pdf>](http://delos-noe.iei.pi.cnr.it/activities/researchforum/Brainstorming/brainstorming-report.pdf)

Digital Library Federation - *A working definition of digital library* [Em linha]. [Consult. 01 Abr. 2009]. Disponível na [www:<URL:http://www.diglib.org/about/dldefinition.htm>](http://www.diglib.org/about/dldefinition.htm)

DIONÍSIO, João – *A Biblioteca Digital da Faculdade de Letras. Em busca da trivialidade* [Em linha]. [Consult. 10 de Jul. 2009]. Disponível na [www:<URL:http://www.goethe.de/ins/pt/lis/pro/bib/dib/dionisio.pdf>](http://www.goethe.de/ins/pt/lis/pro/bib/dib/dionisio.pdf)

DECRETO-LEI n.º 16/2008 – *Direitos de propriedade intelectual* [Em linha]. [Consult. 8 de Fev. 2009]. Disponível na [www:<URL:http://dre.pt/pdf1sdip/2008/04/06400/0189401983.PDF>](http://dre.pt/pdf1sdip/2008/04/06400/0189401983.PDF)

FERREIRA, Miguel – *Introdução à preservação digital: Conceitos, estratégias e actuais consensos* [em linha]. Guimarães: Escola de Engenharia da Universidade do Minho, 2006. [Consult. 12 Ago. 2009]. Disponível na [WWW<URL:https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf>](https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf)

FERRER SAPENA, Antonia [et al.] - *Guía metodológica para la implementación de una biblioteca digital universitaria*. Gijón: Trea, 2005. 202 p. (Biblioteconomía y administración cultural, n.º 113).

FLECKER, Dale – *Preserving Digital Periodicals*. [Em linha], [Consult. 30 Jan 2009]. Disponível na [www:<URL:http://www.digitalpreservation.gov/library/docs/es_periodicals.pdf>](http://www.digitalpreservation.gov/library/docs/es_periodicals.pdf)

FREY, Franziska S.; REILLY, James M. - *Digital Imaging for Photographic Collections: Foundations for Technical Standards* [Em linha]. 2006. [Consult. 25 Set. 2009]. Disponível na [www:<URL: http://www.imagepermanenceinstitute.com/shtml_sub/digibook.pdf>](http://www.imagepermanenceinstitute.com/shtml_sub/digibook.pdf)

Good practice handbook [Em linha]. Ed by Minerva Working Group 6, 2003. [Consult. 27 Set. 2008]. Disponível na [www:<URL: http://www.minervaeurope.org/structure/workinggroups/goodpract/document/bestpracticehandbook1_2.pdf>](http://www.minervaeurope.org/structure/workinggroups/goodpract/document/bestpracticehandbook1_2.pdf)

HUGHES, Lorna M. - *Digitizing collections: strategic issues for the information manager*. London: Facet Publishing, 2004. 288 p.

IFLA – *Directrizes para materiais audiovisuais e multimédia em bibliotecas e outras instituições* [Em linha]. 2006. [Consult. 10 Fev. 2009]. Disponível na [www:<URL: http://archive.ifla.org/VII/s35/pubs/avm-guidelines04-pt.pdf>](http://archive.ifla.org/VII/s35/pubs/avm-guidelines04-pt.pdf)

ISAÍAS, Pedro – *Bibliotecas digitais*. Lisboa: Universidade Aberta, 1999. 176 p.

JUÁREZ SANTAMARÍA, Beatriz; MARTÍNEZ ORTEGA, Patricia - *Los metadatos y la información digital* [Em linha]. 2002. [Consult. 23 Fev. 2009]. Disponível na [www:<URL: http://www.congreso-info.cu/UserFiles/File/Info/Info2002/Ponencias/107.pdf>](http://www.congreso-info.cu/UserFiles/File/Info/Info2002/Ponencias/107.pdf)

LEVY, David M.- *Digital Libraries and the Problem of Purpose* [Em linha]. Jan. 2000. [Consult. 03 Set. 2009]. Disponível na [www:<URL: http://www.asis.org/Bulletin/Aug-00/levy.html>](http://www.asis.org/Bulletin/Aug-00/levy.html)

MARTINEZ EQUIHUA, Saúl – *Biblioteca Digital: conceptos, recursos y estándares*. Buenos Aires: Alfagrama, 2007. 230 p. (Biblioteca Alfagrama).

METS – Metadata Encoding and Transmission Standard [Em linha]. [Consult. 10 Set. 2009]. Disponível na [www:<URL: http://www.loc.gov/standards/mets/>](http://www.loc.gov/standards/mets/)

NISO Framework Working Group - *A Framework of Guidance for Building Good Digital Collections* [Em linha]. 3.^a ed. Baltimore: NISO, 2007. [Consult. 10 Ago. 2008]. Disponível na [www:<URL: http://www.niso.org/publications/rp/framework3.pdf>](http://www.niso.org/publications/rp/framework3.pdf)

NP – 405. 2001, Informação e documentação – *Referências bibliográficas: documentos electrónicos*. Lisboa: IPQ. 28 p.

NP – 405. 1994, Informação e documentação – *Referências bibliográficas: documentos impressos*. Lisboa: IPQ. 49 p.

OCLC – *College Students, Perceptions of Libraries and Information Resources* [Em linha]. Ohio: OCLC, 2007. [Consult. 23 Maio 2009]. Disponível na [www:<URL: http://www.oclc.org/reports/pdfs/studentperceptions.pdf>](http://www.oclc.org/reports/pdfs/studentperceptions.pdf)

PEDROSA, Gilberto – *ContentE: editor de conteúdos* [Em linha]. Biblioteca Nacional Digital. [Consult. 09 de Jul. 2008]. Disponível na [www:<URL: http://bnd.bn.pt/tec/Software/Contente/contente-manual-pt.pdf>](http://bnd.bn.pt/tec/Software/Contente/contente-manual-pt.pdf) (actualmente indisponível)

QUIVY, Raymond; CAMPENHOUDT, Luc Van – *Manual de investigação em ciências sociais*. 4.^a ed. Lisboa: Gradiva, 2005. 284 p.

RAMOS, Sérgio – *Introdução à metodologia do trabalho de projecto* [Em linha]. Fev. 2008 [Consult. 24 de Jul. 2009]. Disponível na [www:<URL: http://livre.fornece.info/media/download_gallery/recursos/metodologia_projecto/TIC-Metodologia-Projecto.pdf>](http://livre.fornece.info/media/download_gallery/recursos/metodologia_projecto/TIC-Metodologia-Projecto.pdf)

Regras portuguesas de catalogação. Lisboa: Instituto Português do Património Cultural. Departamento de Bibliotecas, Arquivos e Serviços de Documentação, 2000. 280 p.

RIBEIRO, Ilda Campos; MENDES, Laurinda – *Biblioteca digital* [Em linha]. [Consult. 13 de Jul. 2009]. Disponível na [www:<URL: http://www.di.ubi.pt/~api/digital_library.pdf>](http://www.di.ubi.pt/~api/digital_library.pdf)

RIEGER, Oya Y.; ENTLICH, Richard; KENNEY, Anne R. - *Moving Theory into Practice: Digital imaging tutorial* [Em linha]. Nova York: Cornell University Library/Research Department, 2003. [Consult. 10 Set. 2009]. Disponível na [www:<URL: http://www.library.cornell.edu/preservation/tutorial/contents.html>](http://www.library.cornell.edu/preservation/tutorial/contents.html)

RILEY, Jenn; WHITSET, Kurt – Features: Practical quality control procedures for digital imaging projects. *OCLC Systems & Services: International Digital Library Perspectives*

[Em linha]. Vol. 21, n.º 1 (2005), p. 40 - 48. [Consult. 10 Set. 2009]. Disponível na
www:<URL: <http://www.dlib.indiana.edu/~jenlrile/publications/imageqc/qc.pdf>>

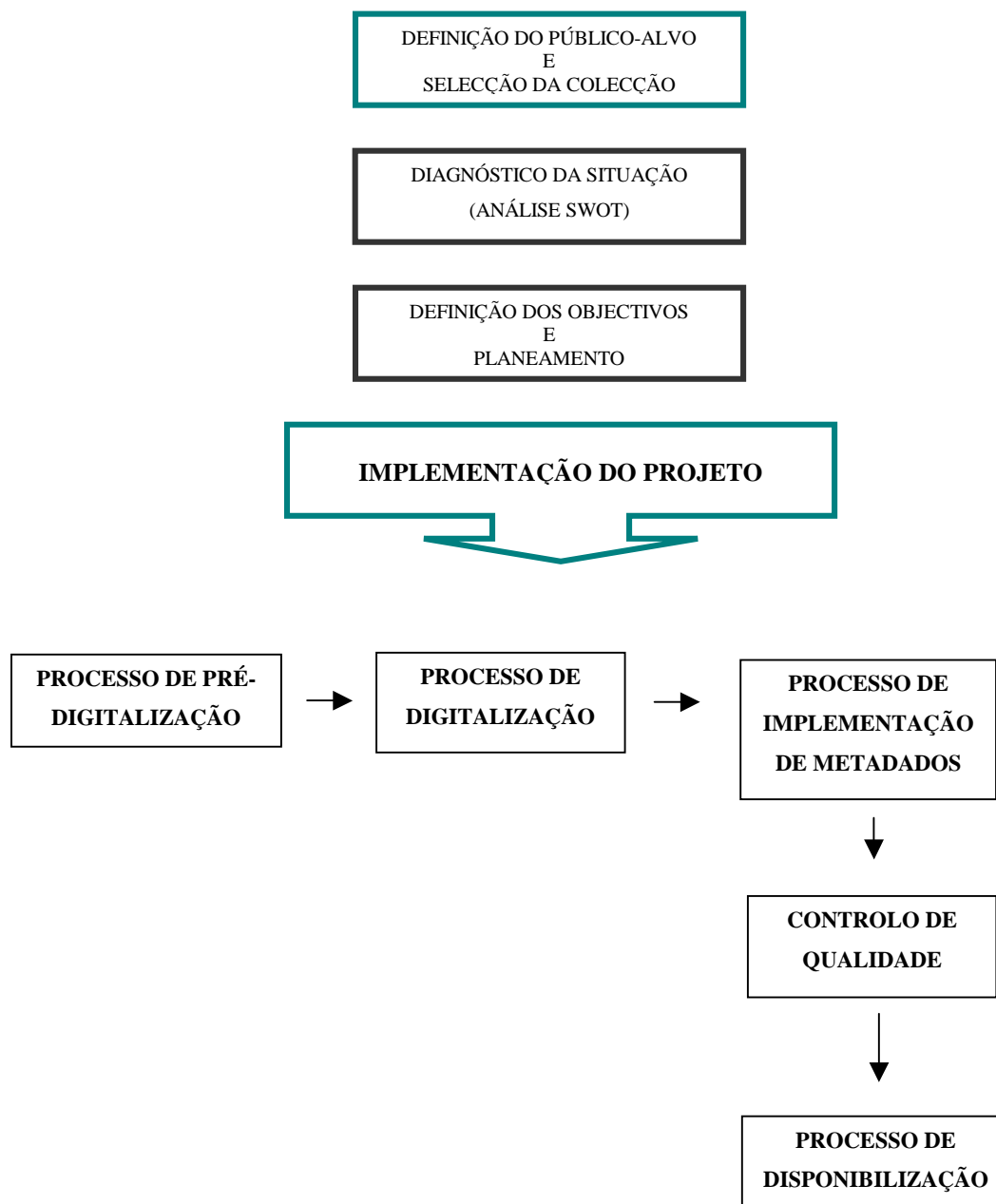
SIBUL – *Sistema Integrado das Bibliotecas da Universidade de Lisboa* [Em linha].
[Consult. 09 de Jul. 2008]. Disponível na
www:<URL:http://aleph18.sibul.ul.pt/F/PISTK928M21XF43EKPGTNNCCHFJRHF14QTUX7IHY11XRJGG5XL-19246?RN=253104637&pds_handle=GUEST>

WITTEN, Ian H; BAINBRIDGE, David - *How to build a digital library*. San Francisco:
Morgan Kaufmann, 2003. 517 p.

ANEXOS

ANEXO 1

CRONOGRAMA DA CONSTITUIÇÃO DO PROJECTO



ANEXO 2

EQUIPAMENTOS TÉCNICOS



Figura 1 – Scanner Kodak i1220 com alimentador



Figura 2 – Scanner Kodak i1220 de mesa ou flatbed

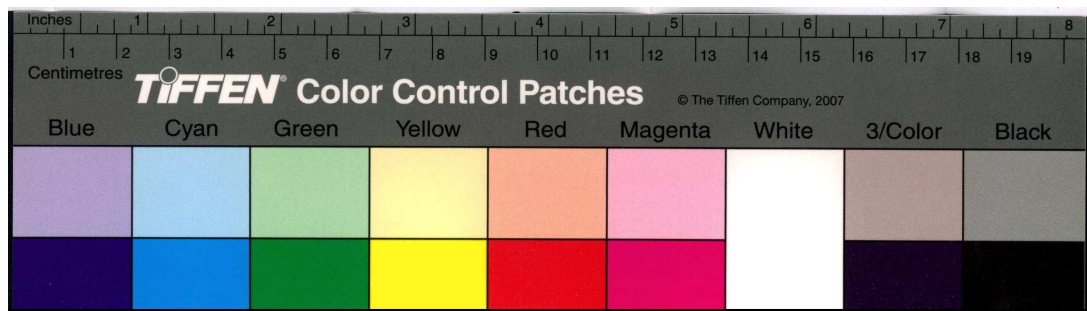


Figura 3 – “Cunha de cor”

Geral	Tipo	Scanner
	Tamanho máximo do meio suportado	215 x 863 mm
	Largura	33 cm
	Profundidade	16.2 cm
	Altura	24.6 cm
	Peso	5.5 kg
	Tipo de entrada	Cor
	Intensidade da Escala de Cinzas (Externo)	8-bit (256 cinza níveis)
	Intensidade de Cor	Cor 48-bit
	Intensidade de Cor (Externo)	24-bit (16,7 milhões de cores)
	Resolução óptica	600 ppp x 600 ppp
	Modo de scanner	Passo único/duplo
	Duplexagem Automática	Sim
Scanner	Tipo de elemento de scanner	CCD
	Tipo de lâmpada/fonte de luz	Lâmpada fluorescente fria
	Velocidade Máxima de Digitalização de Documentos a P&B	30 ppm
	Ciclo de obrigação	1500 varreduras por dia
Manuseamento de media	Detalhes de velocidade do scanner	30 ppm - escala de cinza - 200 dpi 30 ppm - preto & branco - 200 dpi 60 ipm - duplex - escala de cinza - 200 dpi 60 ipm - duplex - preto & branco - 200 dpi
	Padrões de conformidade	TWAIN, ISIS
	Tamanho máximo do documento	215 mm x 863 mm
	Tipo de média suportado	Papel liso
Expansão / conectividade	Tipo de Alimentador de Media	Alimentação automática
	Capacidade de alimentador	50 folhas
	Interfaces	1 x Hi-Speed USB - 4 PIN USB Tipo B
Diversos	Padrões de conformidade	CE, UL, TUV GS, BSMI, VCCI Class B ITE, CISPR 22 Class B, BSMI CNS 13438 Class B, cUL, EN 60950, IEC 60950, EN 61000-3-3, EN55022 Class B, ICES-003 Class B, C-Tick Class B, FCC Part 15 B, CCC, AS/NZS 3548, EN55024 Class B, FCC CFR47 Part 15, UL 60950-1, CSA C22.2 No. 60950-1-03
Software / Requisitos do sistema	Software incluído	Drivers & Utilities, ScanSoft OmniPage, ScanSoft PaperPort, KODAK Capture Software Lite
	SO necessário	Microsoft Windows 2000 / XP

Figura 4 - Características do Equipamento Kodak i1220

ANEXO 3

PROCESSO DE LEVANTAMENTO

O processo de levantamento das revistas, alvo de digitalização, possui várias etapas que de seguida apresentaremos.

1) Pesquisar no catálogo colectivo SIBUL as existências da revista a tratar. Pesquisa-se pelo catálogo Faculdade de Letras, seguido de pesquisa avançada, procura-se pelo título e assinala-se o mesmo como um periódico.

The screenshot shows the SIBUL search interface. At the top, there's a header with the SIBUL logo and navigation links. The main section is titled 'Pesquisa Avançada (Fac. Letras)'. It contains a search form with fields for 'Campo a pesquisar' (Title, Palavra (Geral), Palavra (Geral)), 'Digite palavra ou frase' (with the example 'artis'), and 'Palavras adjacentes?' (radio buttons for Não and Sim). There are also buttons for 'Pesquisar' and 'Limpar'. Below the search form, there's a section for 'Limitar pesquisa a:' with filters for 'Lingua' (Todas as línguas), 'Formato' (Periódicos), 'Ano (de .. a)' (with a truncation option), and 'Cota' (Todas as bibliotecas). At the bottom, there's a section for 'Sugestões para pesquisa' and a note about the search results.

Figura 1

2) Verificação do número de existências da revista que vamos tratar e seu estado actual.

The screenshot shows the search results page for the journal 'História da arte'. It displays the journal's details, including its ISSN (000366854), description, and CDU classification (7(091)(05)). Below the details, there's a section for 'O que pode fazer a partir dos exemplares apresentados:' with instructions on how to request a reserve, requisition, or photocopy. The main part of the page is a table showing the list of exemplars, including their description, state, date of return, and location. The table has columns for 'Ano de publicação', 'Descrição', 'Estado do exemplar', 'Data de devolução', 'Hora de devolução', 'Biblioteca', 'Cota', 'Cota-2', 'Cód. de barras', 'Nota para o público', and 'SFX'. The table lists 7 exemplars, all from the 'ULFL' library, with various issue numbers and dates.

Ano de publicação	Descrição	Estado do exemplar	Data de devolução	Hora de devolução	Biblioteca	Cota	Cota-2	Cód. de barras	Nota para o público	SFX
Nº 7 (Dez. 2008)	Não requisitável	Esperado em 30/01/09			ULFL	7(05) ART	PP 731	55172-70		SFX
Nº 6 (Dez. 2007)	Não requisitável	Na estante			ULFL	7(05) ART	PP 731	55172-60		SFX
Nº 5 (Dez. 2006)	Leitura na Sala	Na estante			ULFL	7(05) ART	PP 731	55172-50		SFX
Nº 4 (Dez. 2005)	Não requisitável	Na estante			ULFL	7(05) ART	PP 731	55172-40		SFX
Nº 3 (Dez. 2004)	Não requisitável	Na estante			ULFL	7(05) ART	PP 731	55172-30		SFX
Nº 2 (Out. 2003)	Não requisitável	Na estante			ULFL	7(05) ART	PP 731	55172-20		SFX
Nº 1 (Out. 2002)	Não requisitável	Na estante			ULFL	7(05) ART	PP 731	55172-10		SFX

Figura 2

- 3) Assinalar os exemplares existentes, verificando se existem números em duplicado e triplicado.
- 4) Registrar todas as informações no ficheiro Excel denominado “Listagem de Títulos” preenchendo os campos *Ano*; *Tomo*; *Série*; *Número*; *Código de barras*; *Duplicado*; *Triplicado*; *Observações* e *Estado* na *sheet* da respectiva revista. Como no exemplo seguinte:

ARTIS								
Ano	Tomo	Série	Número	Código de barras	Duplicado	Triplicado	Observações	Estado
2002			1	ULFL55172-10		X		Encontra-se no cesto p guilhotinar
2003			2	ULFL55172-20		X		Encontra-se no cesto p guilhotinar
2004			3	ULFL55172-30		X		Encontra-se no cesto p guilhotinar
2005			4	ULFL55172-40		X		Encontra-se no cesto p guilhotinar
2006			5	ULFL55172-50		X		Encontra-se no cesto p guilhotinar
2007			6	ULFL55172-60		X		Encontra-se no cesto p guilhotinar
2008			7	ULFL55172-70			Exemplar inexistente	

Figura 3 – Ficheiro “Listagem de Títulos” – Revista Artis

Como podemos verificar na fig. 3, a revista Artis tem todos os exemplares em triplicado à excepção do último que tem a nota de “exemplar inexistente”, isto porque é um exemplar que ainda não deu entrada na biblioteca. Contudo, temos casos em que a situação é mais complicada. Observemos o exemplo seguinte:

PHILOSOPHICA								
Ano	Tomo	Série	N.º	Código de barras	Duplicado	Triplicado	Observações	Estado
Abr. 1993			1	ULFL14278-10		X		Encontra-se no cesto p Guilhotinar
Nov. 1993			2	ULFL14278-20		X		Encontra-se no cesto p Guilhotinar
Abr. 1994			3	ULFL14278-30		X		Encontra-se no cesto p Guilhotinar
Nov. 1994			4	ULFL14278-40		X		Encontra-se no cesto p Guilhotinar
Abr. 1995			5	ULFL14278-50		X		Encontra-se no cesto p Guilhotinar
Nov. 1995			6	ULFL14278-60	X			Encontra-se no cesto p Guilhotinar
Abr. 1996			7	ULFL14278-70	X			Encontra-se no cesto p Guilhotinar
Nov. 1996			8	ULFL14278-80	X			Encontra-se no cesto p Guilhotinar
Abr. 1997			9	ULFL14278-90			Exemplar único	
Nov. 1997			10	ULFL14278-100	X			Encontra-se no cesto p Guilhotinar
Abr. 1998			11	ULFL14278-110		X		Encontra-se no cesto p Guilhotinar
Nov. 1998			12	ULFL14278-120	X			Encontra-se no cesto p Guilhotinar
Abr. 1999			13	ULFL14278-130			Exemplar único	
Nov. 1999			14	ULFL14278-140			Exemplar único	
Abr. 2000			15	ULFL14278-150	X			
Nov. 2000			16	ULFL14278-160	X			Encontra-se no cesto p Guilhotinar

Figura 4 – Ficheiro “Listagem de Títulos” – Revista Philosophica

A	B	C	D	E	F	G	H	I	J	L
19	Nov. 2001			17/18	ULFL14278-180	X			Encontra-se no cesto p Guilhotinar	
20	Abr./Nov. 2002			19/20	ULFL14278-190	X			Encontra-se no cesto p Guilhotinar	
21	Abr. 2003			21	ULFL14278-200			Exemplar único		
22	Nov. 2003			22	ULFL14278-210	X			Encontra-se no cesto p Guilhotinar	
23	Abr. 2004			23	ULFL14278-220	X			Encontra-se no cesto p Guilhotinar	
24	Nov. 2004			24	ULFL14278-230			Inexistência de exemplar		
25	Abr. 2005			25	ULFL14278-240			Exemplar único		
26	Nov. 2005			26	ULFL14278-250			Exemplar único		
27	Abr. 2006			27	ULFL14278-260			Exemplar único		
28	Nov. 2006			28	ULFL14278-270			Inexistência de exemplar (Nota sibul: Esperado em 30/12/2006)		
29	Abr. 2007			29	ULFL14278-280			Exemplar único		
30	Nov. 2007			30	ULFL14278-290			Inexistência de exemplar		
31	Abr. 2008			31						

Figura 4 (cont.) – Ficheiro “Listagem de Títulos” – Revista Philosophica

Ao fazer o levantamento da revista *Philosophica* deparámo-nos com um problema que teria de ser resolvido antes de se proceder à “guilhotinagem” ou digitalização de qualquer exemplar. Seria necessário encontrar outras cópias para as revistas que são dadas como “exemplares únicos”, no sentido de evitar guilhotinar um exemplar único. Se não existisse na biblioteca, feríamos de solicitar aos departamentos da faculdade que editam as revistas, o número em falta.

Este levantamento foi feito para todas as publicações em série publicadas pelos Departamentos e Centros de Documentação da Faculdade de Letras e registado num ficheiro Excel com todos os dados.

5	2003	2	1	ULFL54664-30	X		
6			2	ULFL54664-40	X		
7	2004	3	1	ULFL54664-50	X		
8			3	2	ULFL54664-60		Inexistência de exemplar
9	2005	4	1	ULFL54664-70			Inexistência de exemplar
10			4	2	ULFL54664-80		Inexistência de exemplar: Nota Sibul: Esperado em 30/12/2006

Figura 6

- 5) Após o levantamento das revistas e respectivo registo no ficheiro, enviam-se os títulos que cumprem as condições estipuladas para a “guilhotinagem”.
- 6) Quando os títulos regressam, já guilhotinados, são verificados um a um para perceber se os códigos de barras e toda a informação coincidem e se não existem trocas. Por fim, colocam-se numa estante como títulos prontos para a digitalização.

ANEXO 4

PROCESSO DE DIGITALIZAÇÃO

Etapas para a constituição de um processo de digitalização:

- 1) Abrir o programa ScandAll e seleccionar *scan to view* para iniciar o processo de digitalização

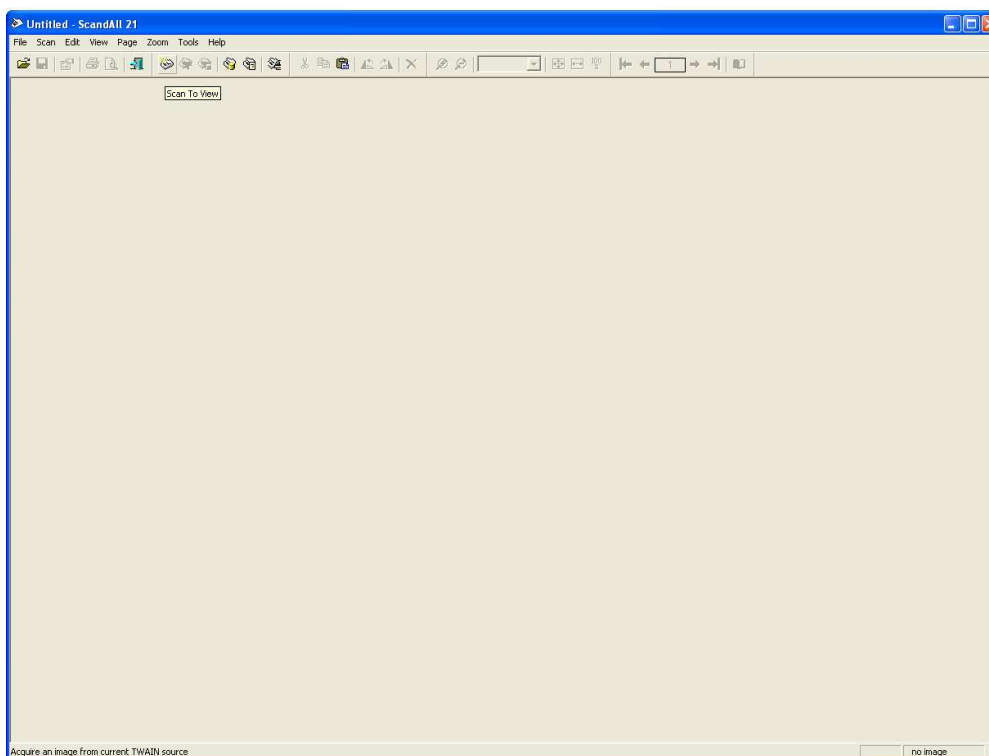


Figura 1

Definiram-se dois modelos devidamente parametrizados, um para a digitalização de capas e contracapas no *flatbed*, outro para a digitalização automática das páginas da revista. Como podemos visualizar no exemplo seguinte, as imagens foram digitalizadas com a profundidade de cor de 24 bits e com a resolução de 300 dpi.

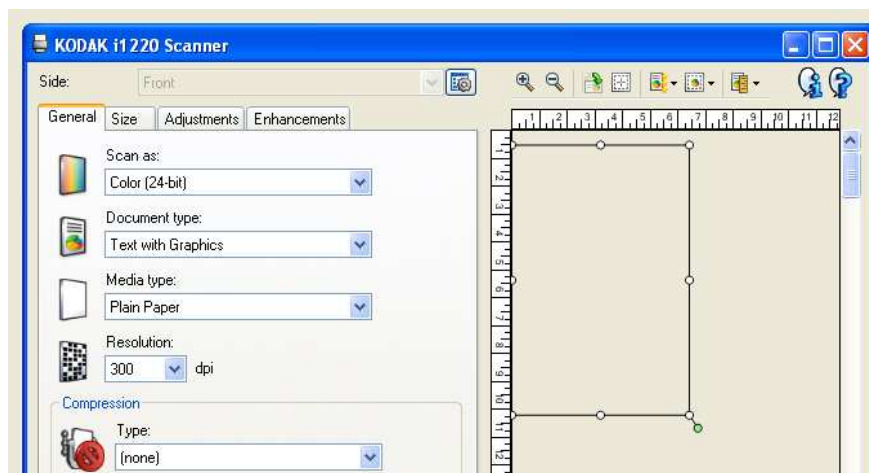


Figura 2

2) Escolher o scanner a utilizar para executar a digitalização.

a) Para a **digitalização das capas** usar o modelo *bib digital flatbed*

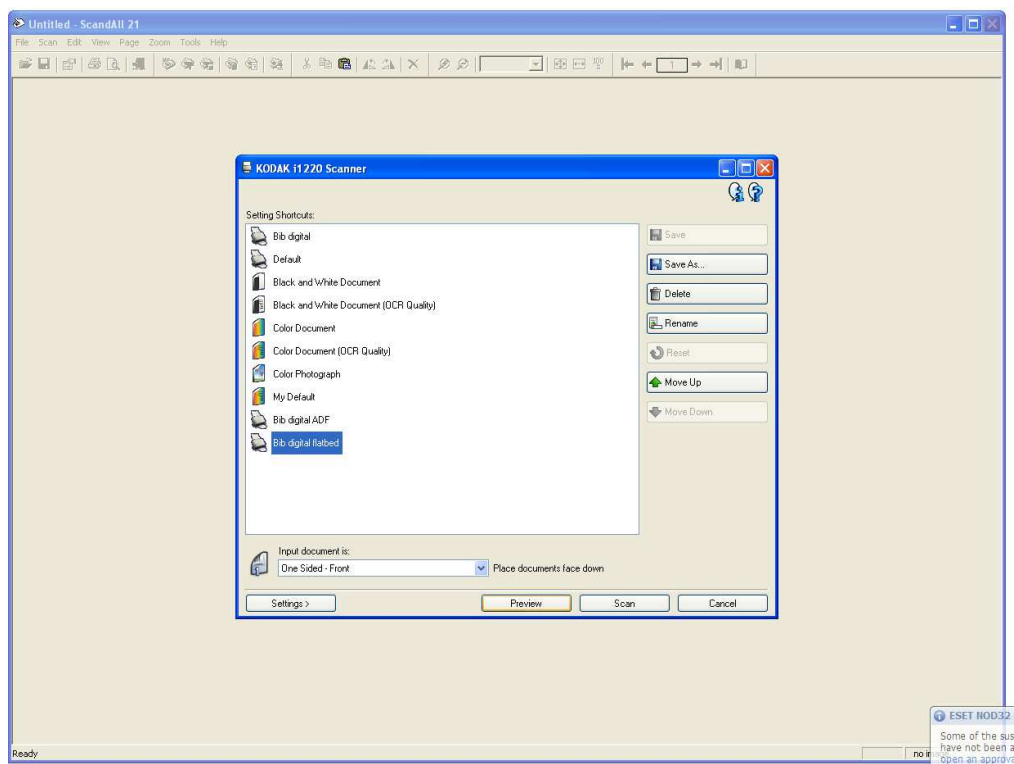


Figura 3

Definir as margens da capa e digitalizar através da função *preview*. De notar que a “régua de cor” e a capa são digitalizadas conjuntamente.

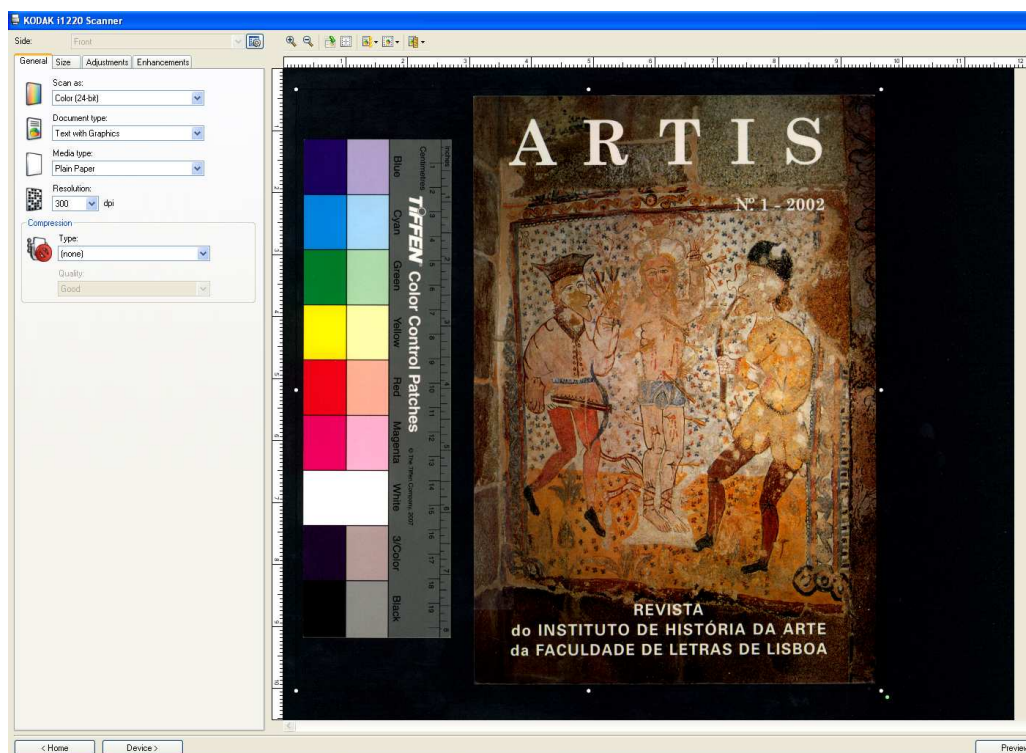


Figura 4

- Digitalizar a capa (scan):

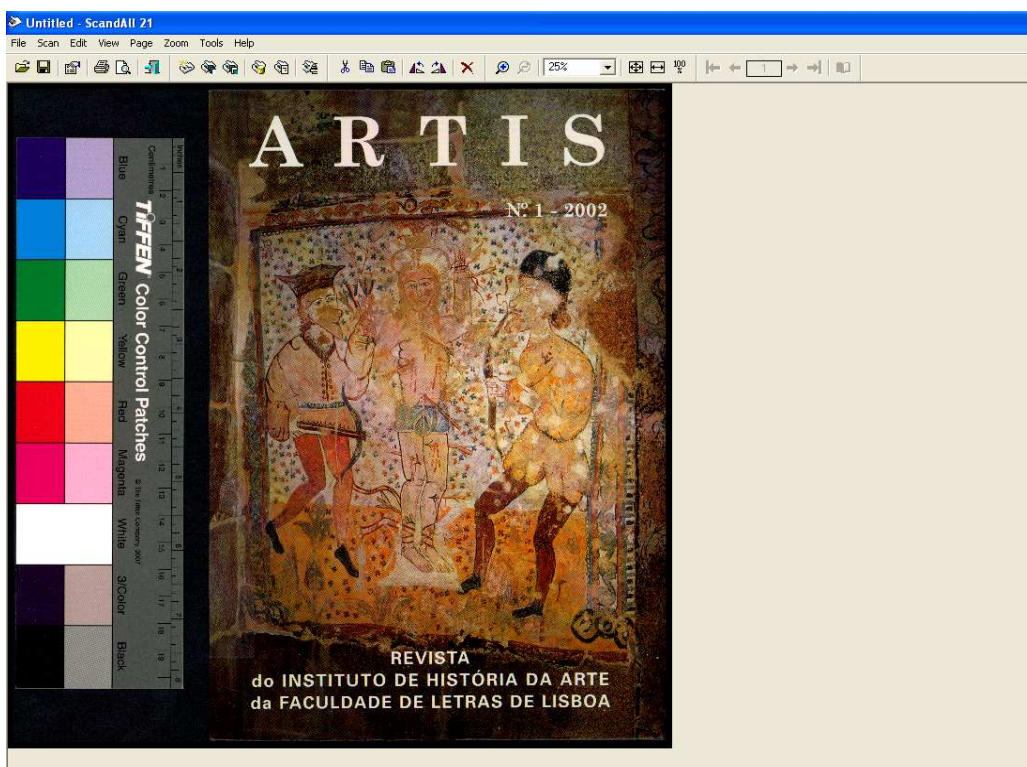


Figura 5

- Guardar a imagem sempre em formato TIFF na pasta correspondente à revista

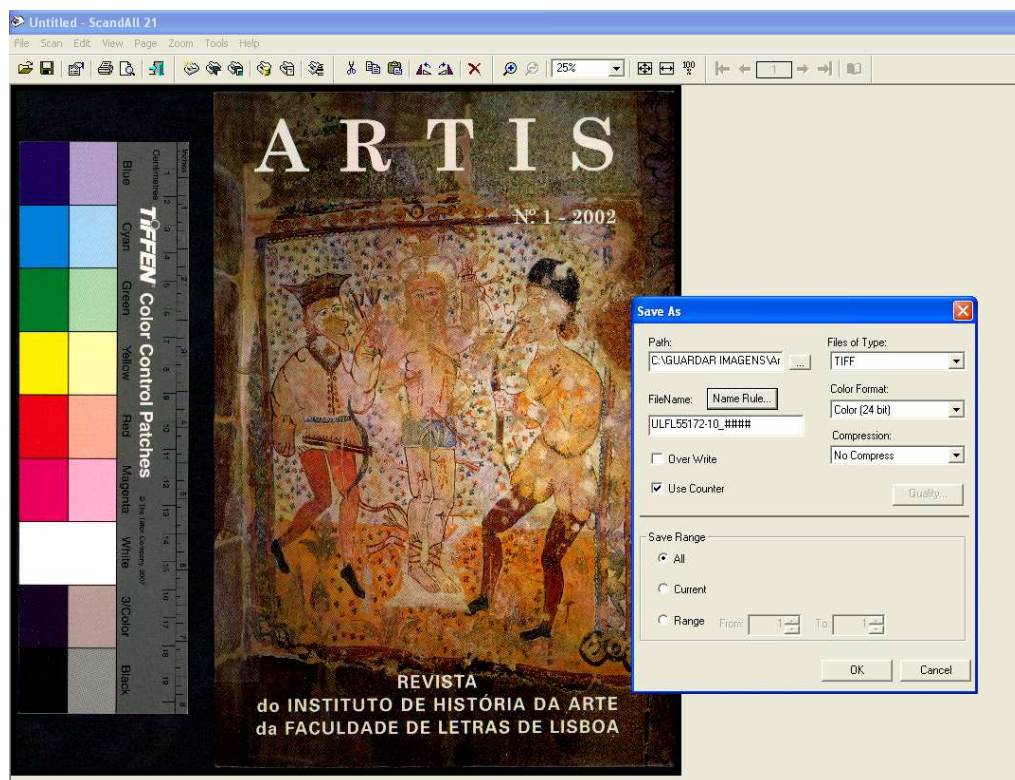


Figura 6

b) Para a **digitalização do conteúdo da obra** usar o modelo *Bib digital ADF*

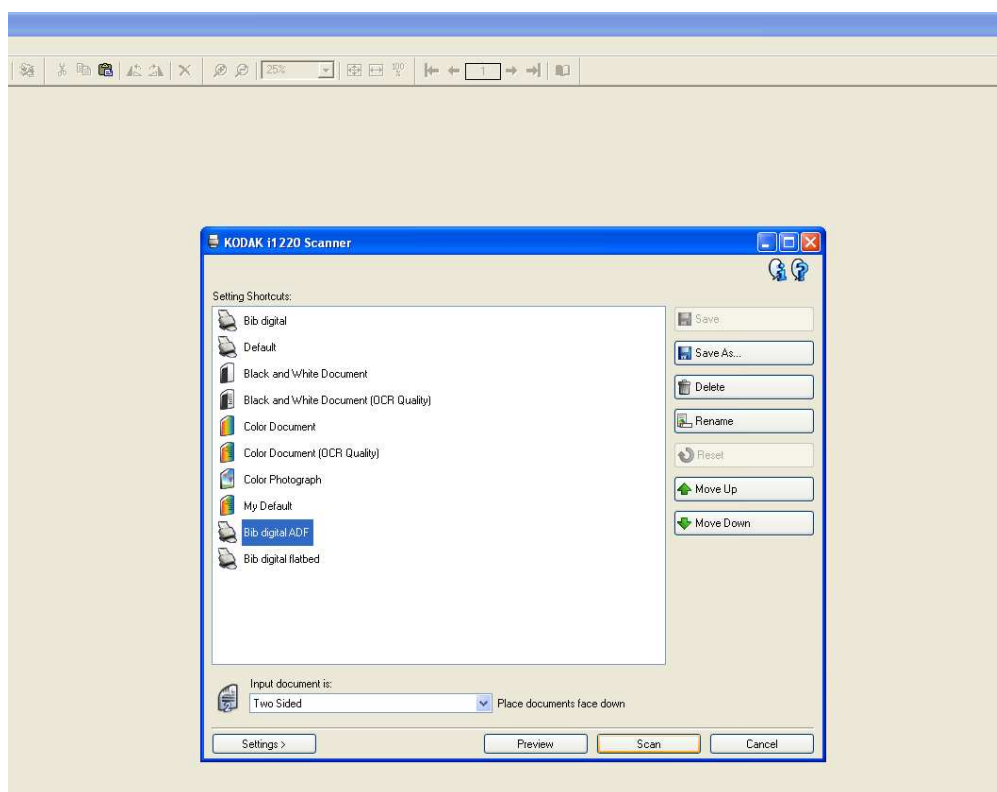


Figura 7

- Digitalizar as páginas através do alimentador automático.

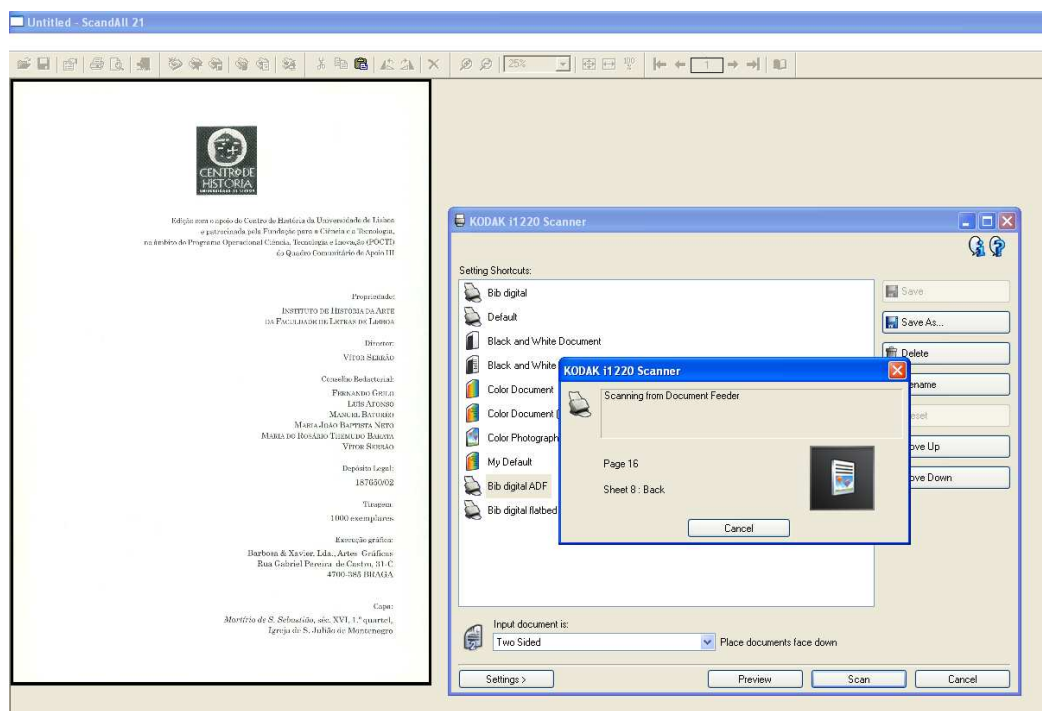


Figura 8

- Guardar estas imagens no caminho: C:\GUARDAR IMAGENS\Artis\ULFL55172-10, ou seja, na pasta da respectiva revista:

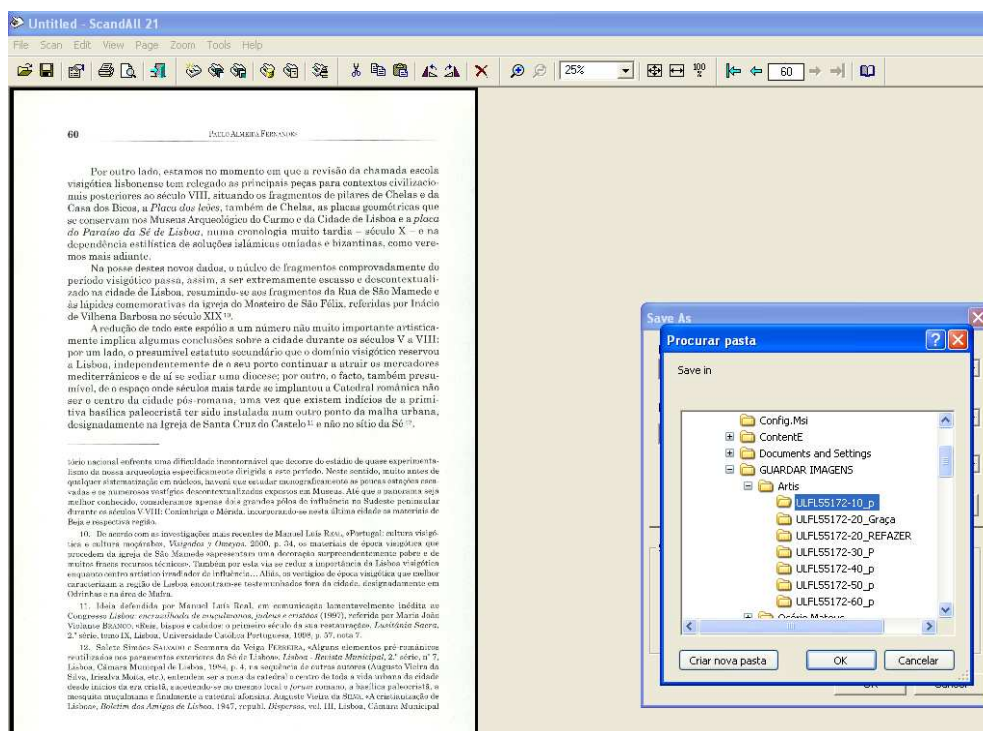


Figura 9

- Guardar as imagens sempre no formato TIFF, a 24 bits e sem compressão.

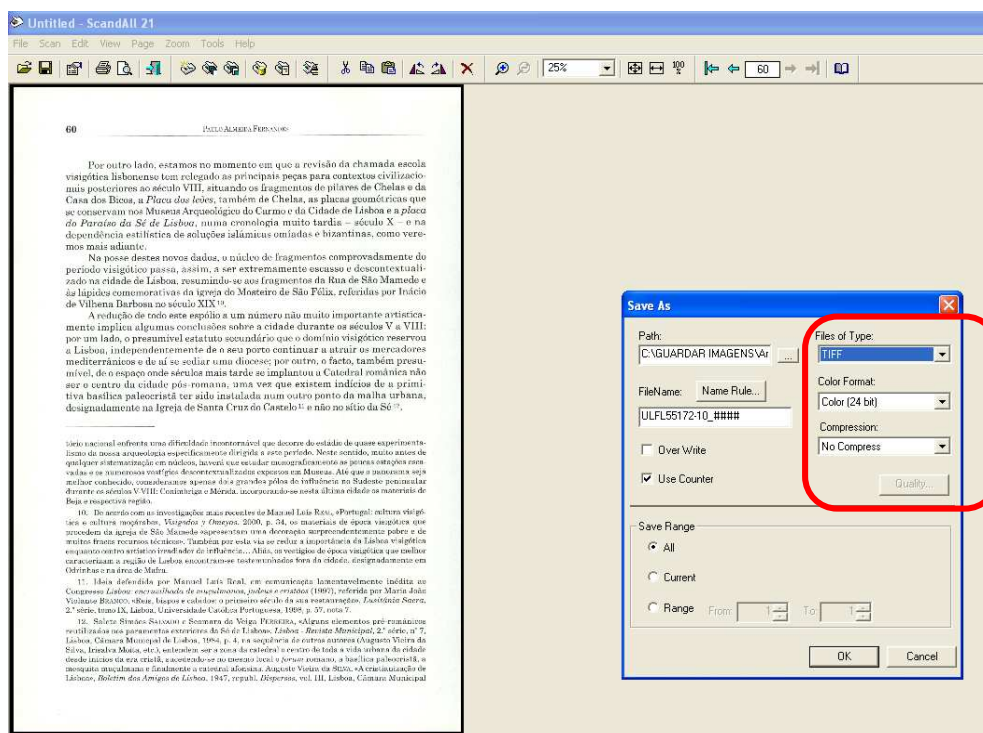


Figura 10

Após a digitalização de toda a obra, e da sua gravação na pasta correspondente à revista, o resultado é um conjunto de imagens denominadas pelo código de barras da revista, seguido do n.º sequencial da imagem que respeitam a ordem do documento original, como podemos verificar no exemplo seguinte:

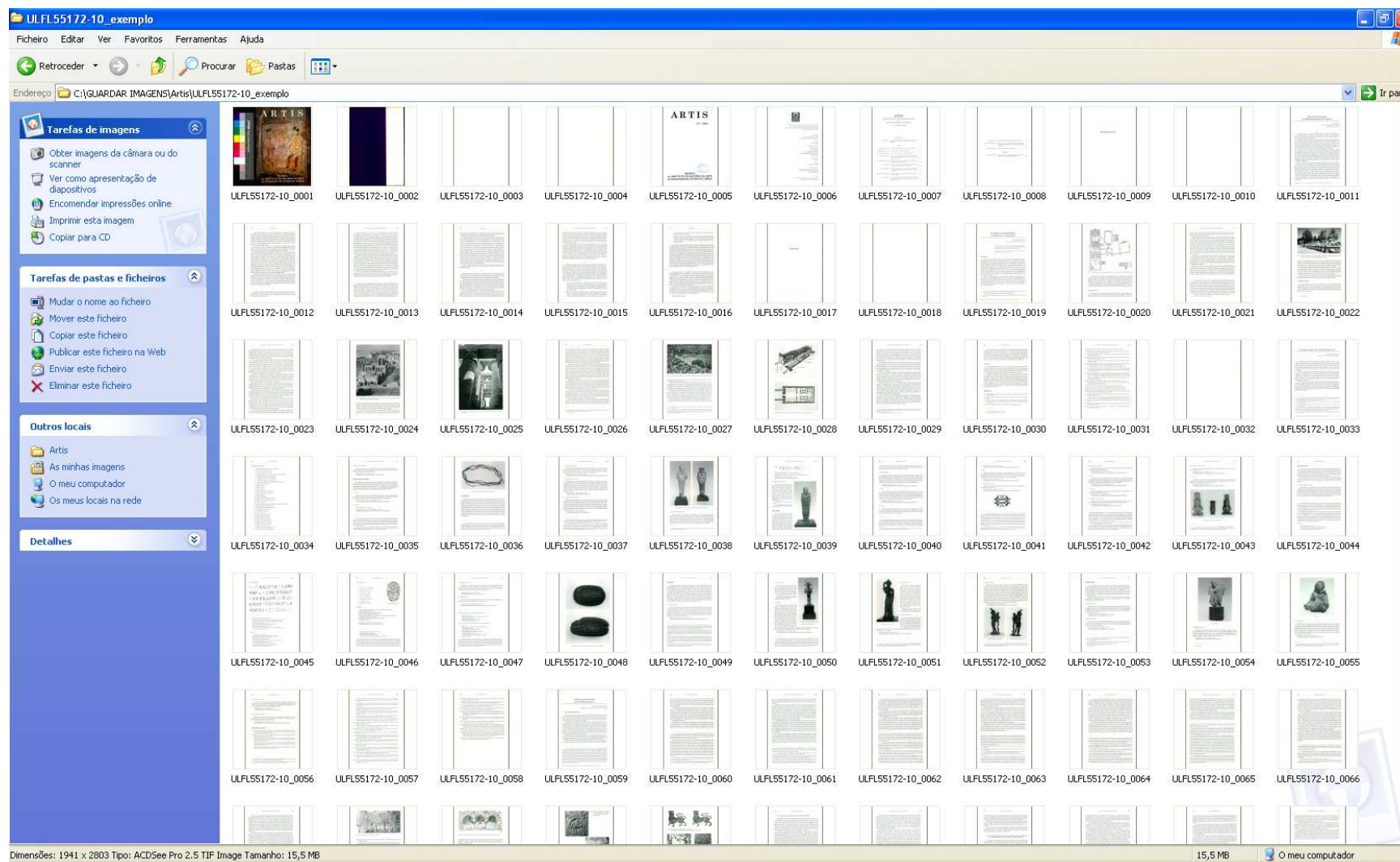


Figura 12 – Resultado final do processo de digitalização

ANEXO 5

PROCESSO DE IMPLEMENTAÇÃO DE METADADOS E NOMEAÇÃO DAS IMAGENS

1) O primeiro passo é abrir o programa PAPAIA e importar as imagens digitalizadas da obra que queremos estruturar e renomear (Fig. 1 e 2)

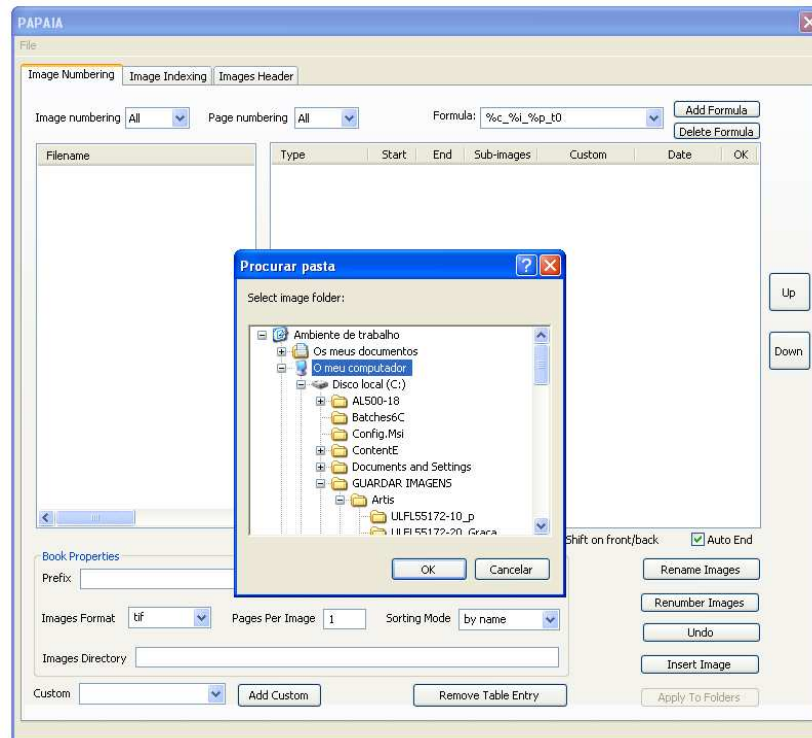


Figura 1

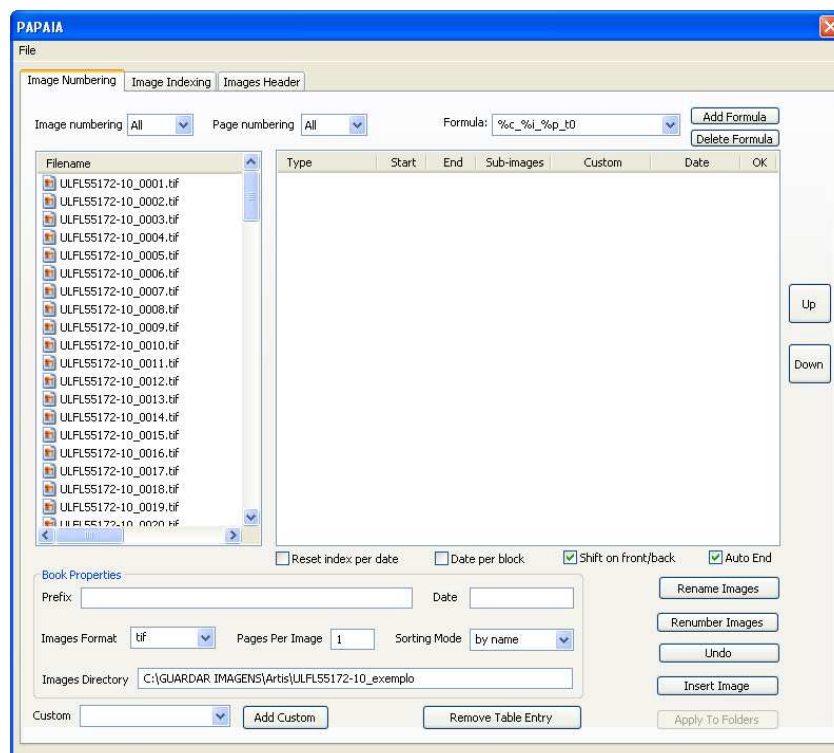


Figura 2

2) De seguida, no separador *Images Header* escolhe-se as *tags* TIFF que irão descrever as imagens, como podemos verificar nos exemplos seguintes (fig. 3 e 4):

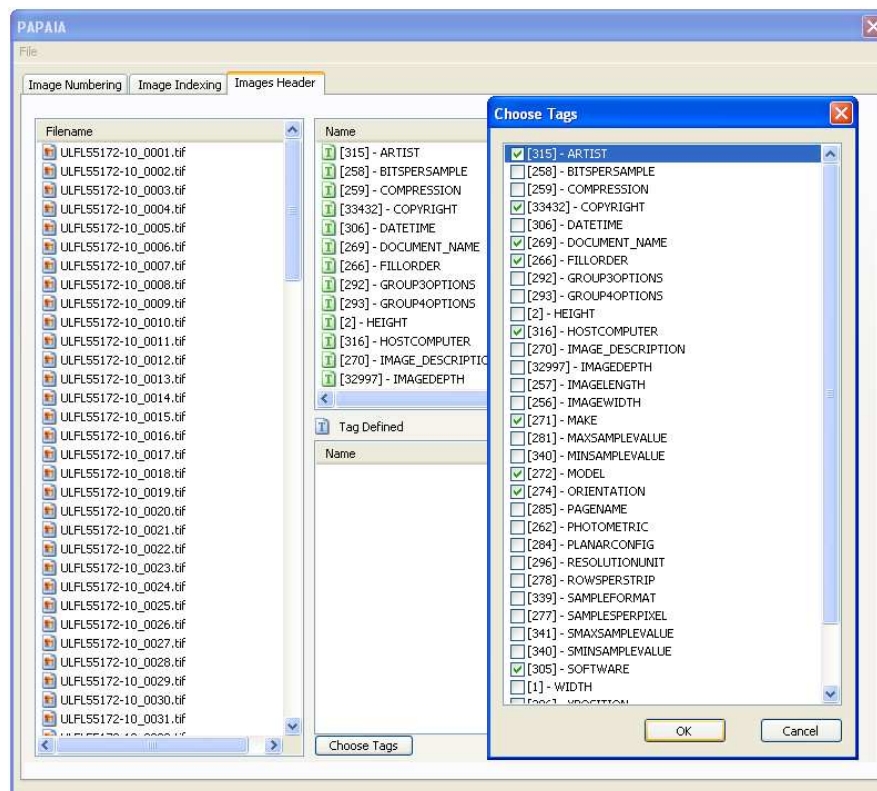


Figura 3

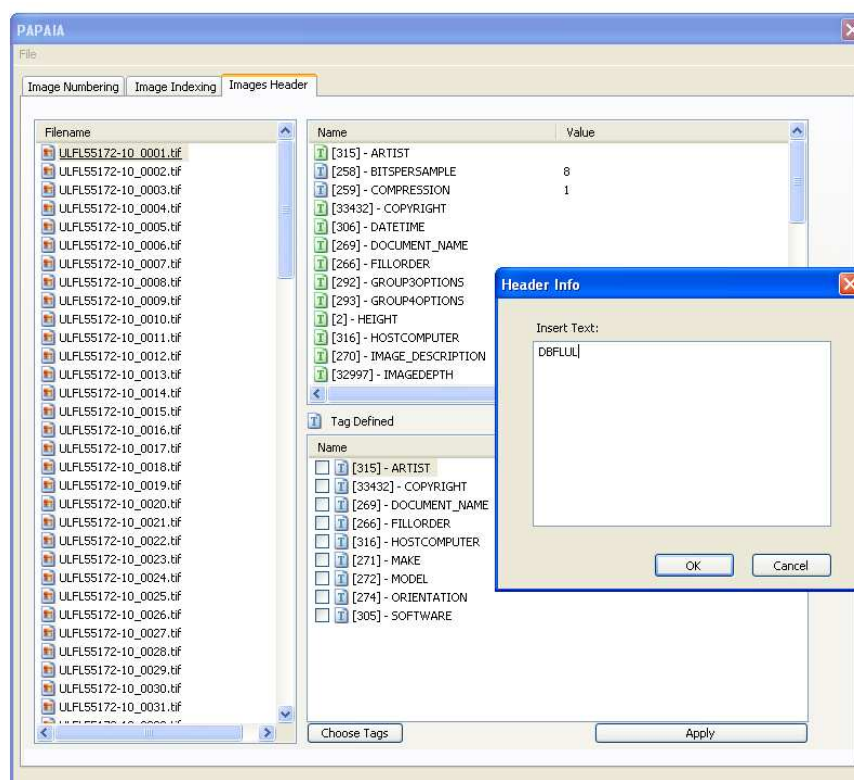


Figura 4

3) Depois de preenchermos manualmente todos os campos, aplicamos os dados a todas as imagens.

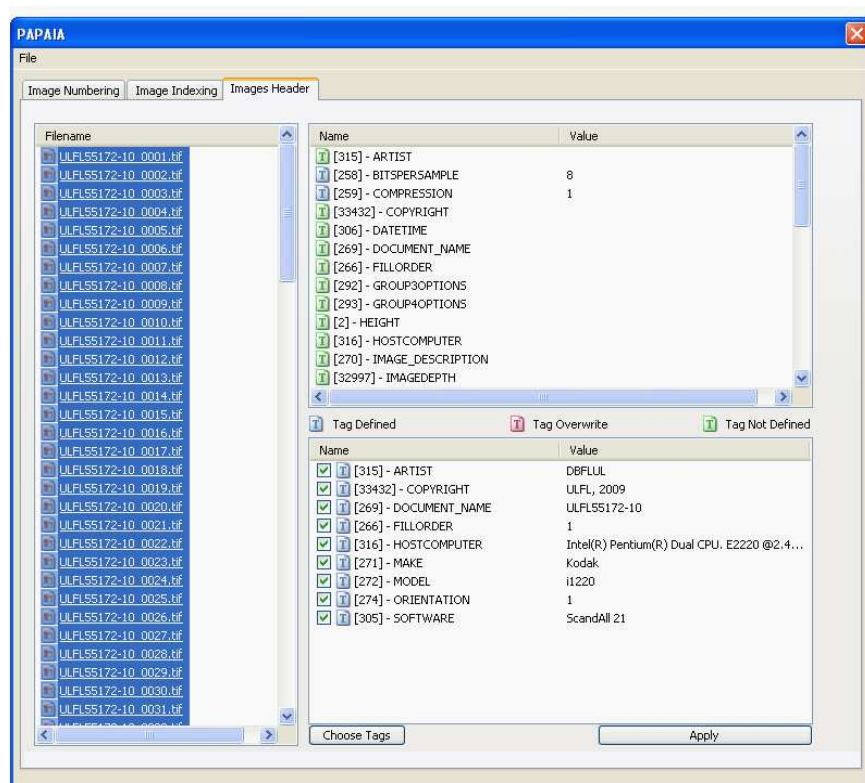


Figura 5

Após a aplicação dos cabeçalhos TIFF e ao seleccionar uma imagem verificamos que os campos que descrevemos já se encontram inseridos na imagem.

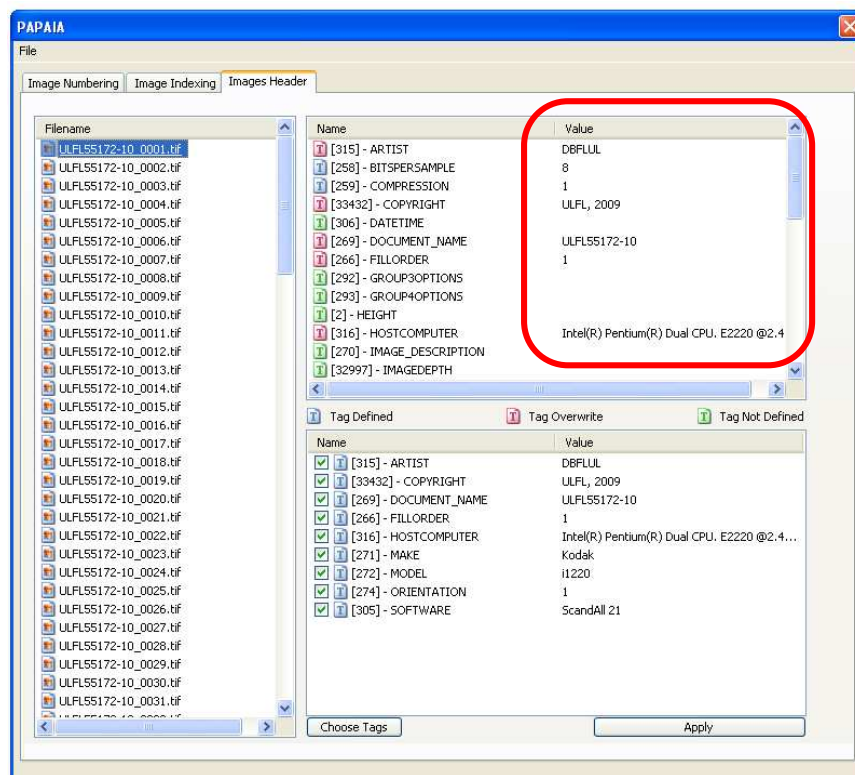


Figura 6

4) No separador *image numbering* configuramos os metadados essenciais para a nomeação correcta dos ficheiros de imagens digitalizadas, sublinhados a vermelho, no exemplo seguinte:

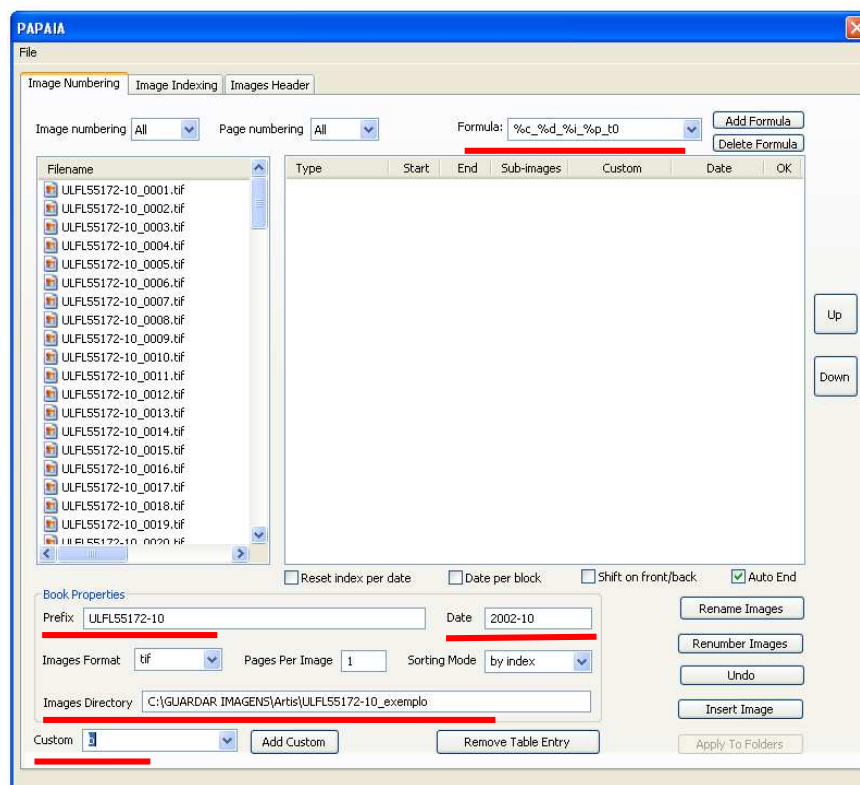


Figura 7

5) Seguidamente, procedemos à **renomeação** de ficheiros propriamente dita através do método “*drag and drop*” das imagens.

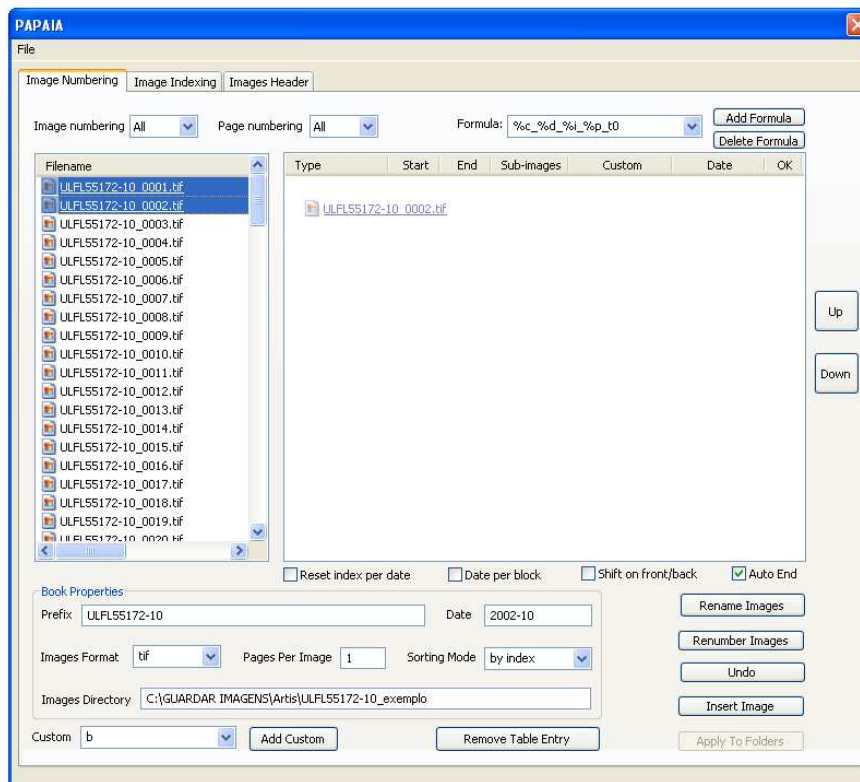


Figura 8

- Através deste programa podemos nomear uma imagem como “capa”, imagem “branca” (b), “rosto” e atribuir-lhes o número sequencial de paginação da obra.

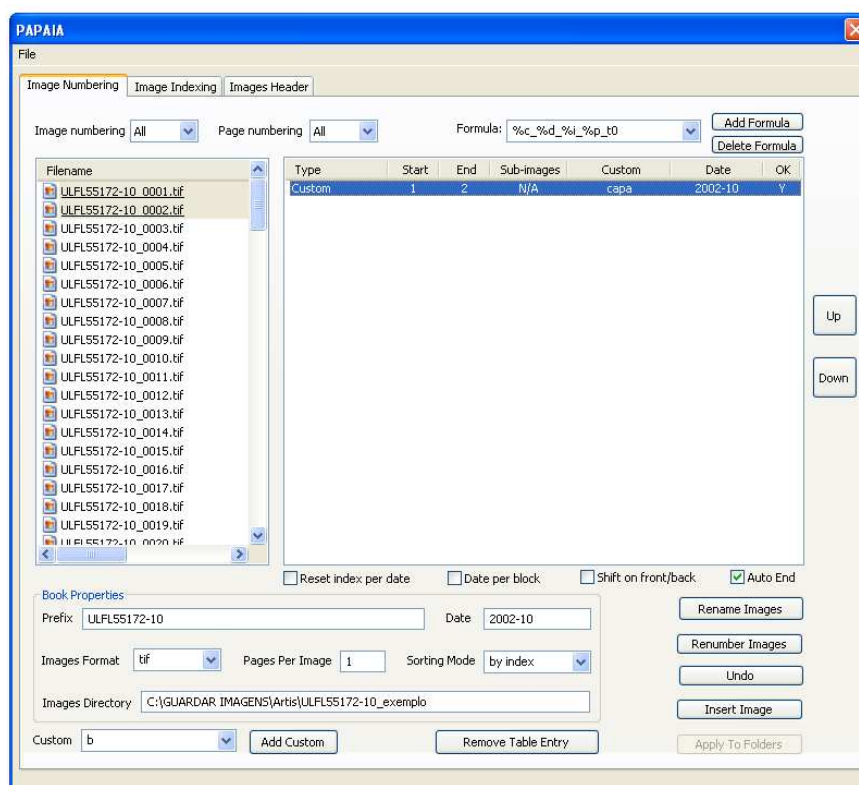


Figura 9

- Realizamos este procedimento para todas as imagens, sendo que podemos nomear conjuntos de imagens como podemos visualizar no exemplo seguinte: 2 imagens capa e 2 imagens brancas:

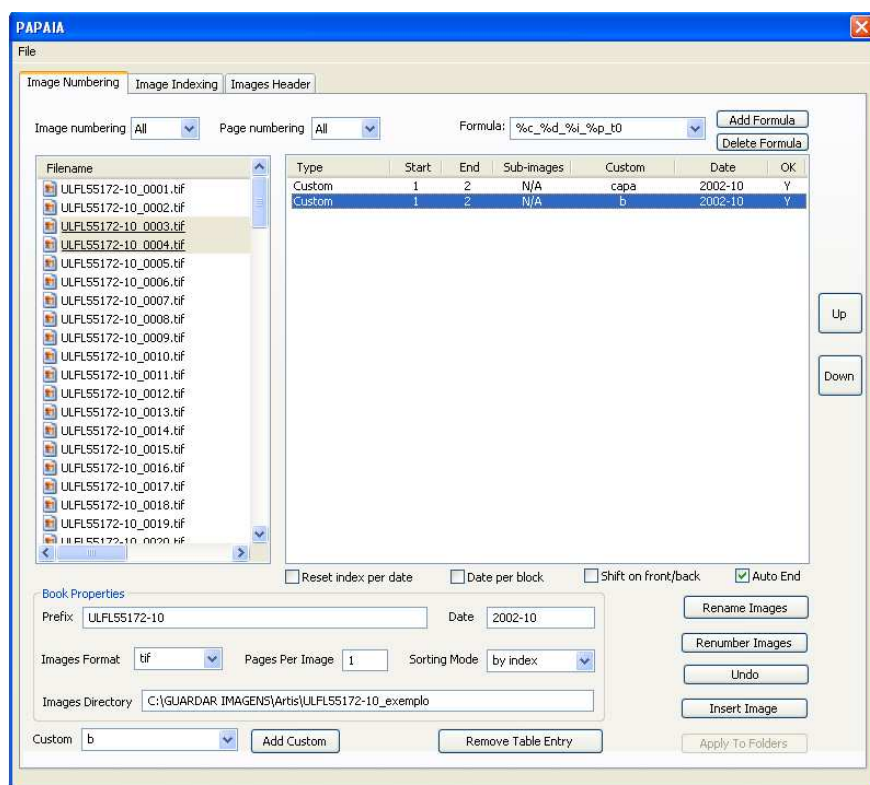


Figura 10

- Os ficheiros designados com o tipo *Custom* têm sempre uma descrição identificativa (capa, b, rosto,...) enquanto que os ficheiros do tipo *Arabic* têm uma descrição numérica, isto porque designam as imagens que têm texto impresso e atribuem o número de paginação que consta na obra. Como podemos verificar no seguinte exemplo, a imagem ULFL55172-10_0090 foi correctamente denominada de b (branca):

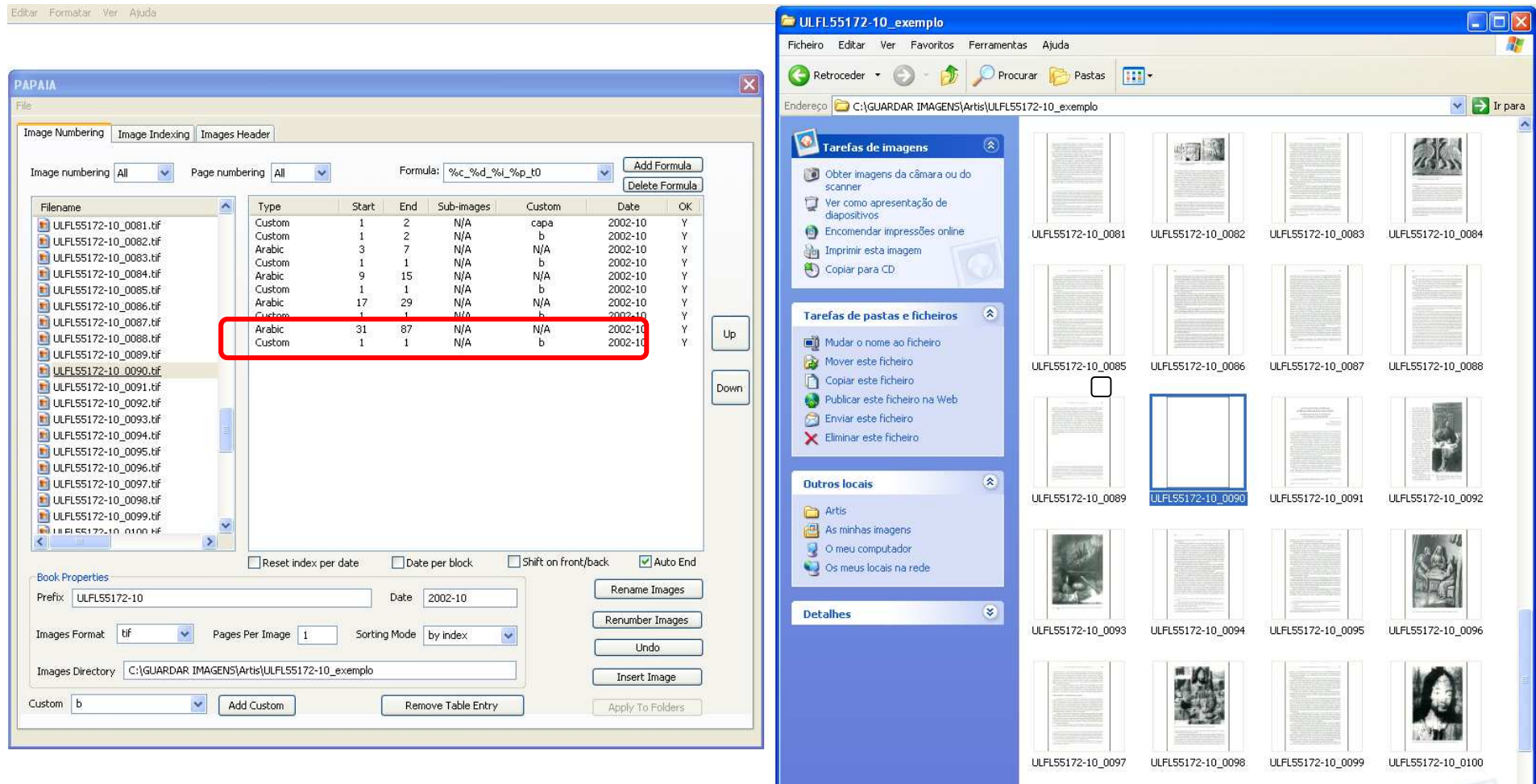


Figura 11

- A paginação atribuída pelo PAPAIA tem de coincidir com o número de página real da revista, como podemos verificar no exemplo seguinte:

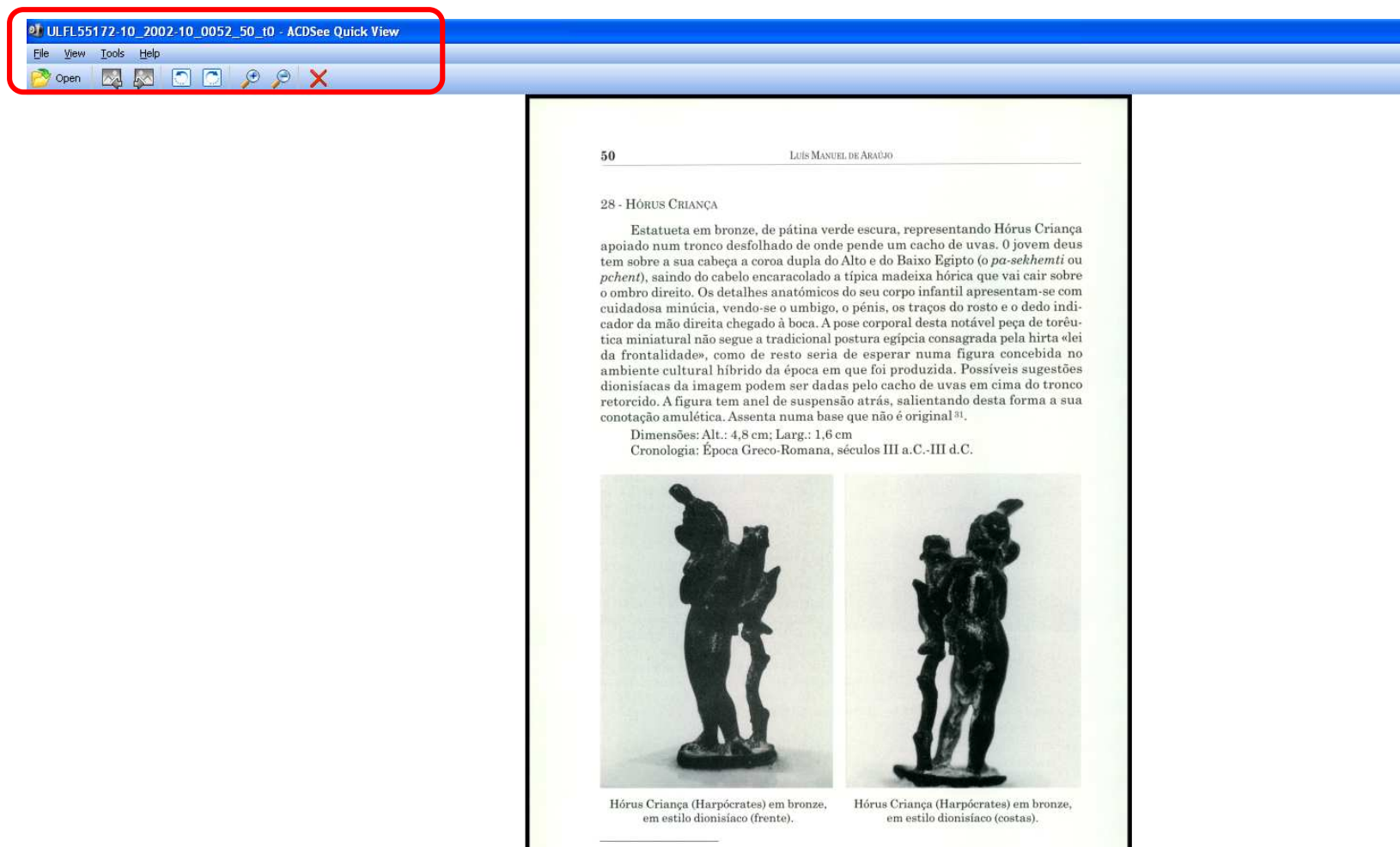


Figura 12

6) Depois de se concretizar a nomeação de todas as imagens, gravamos as alterações efectuadas seleccionando o botão *Rename images*. Como podemos verificar nos exemplos seguintes, as imagens adquirem um novo nome, designadamente: <código_de_barras> _ <data e mês da publicação> – <ordem_da_imagem> _<referência_da_página> _t0, e são guardadas na sua pasta correspondente.

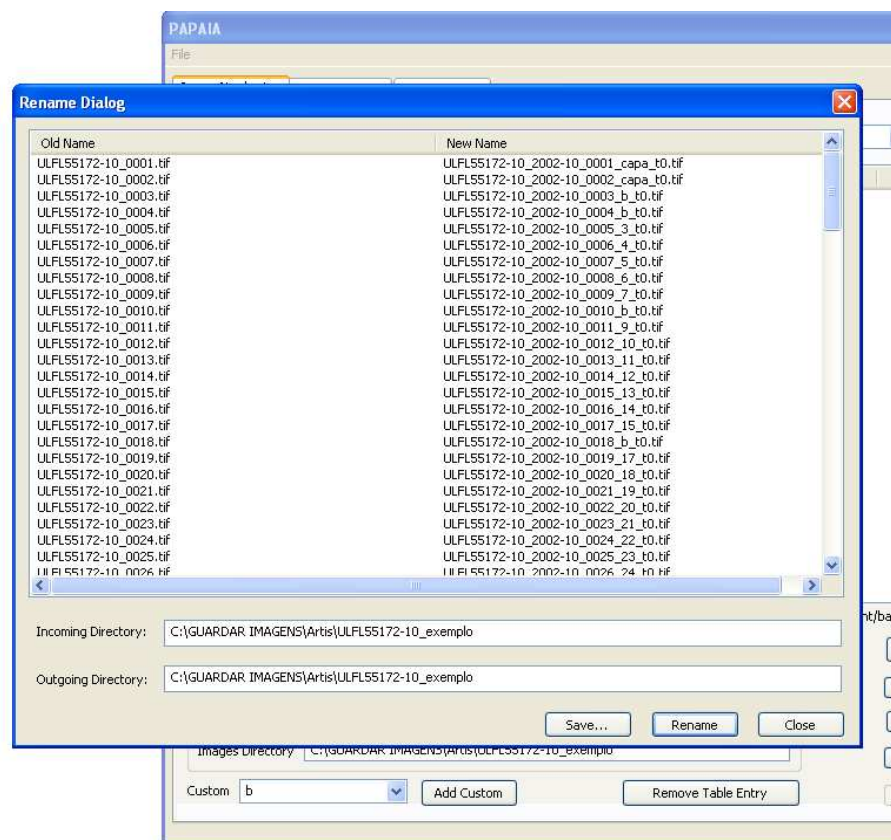


Figura 13

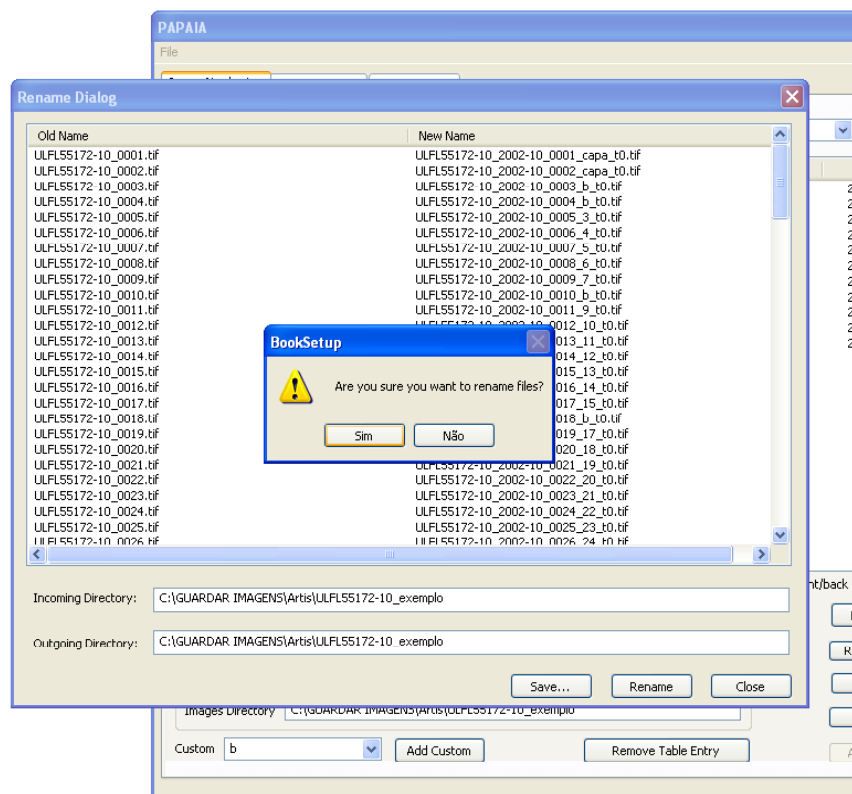


Figura 14

7) Ao guardarmos as imagens a aplicação solicita-nos imediatamente que gravemos o mesmo conteúdo num ficheiro XML. Guardamos o ficheiro numa pasta devidamente identificada.

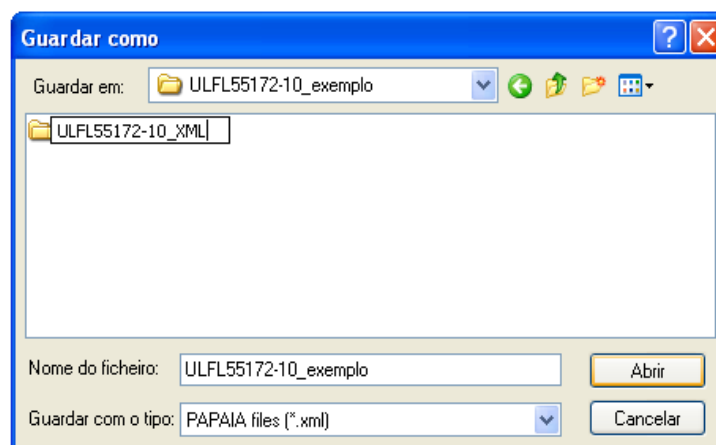


Figura 15

Exemplo de uma estrutura descritiva XML de uma obra exportada pela aplicação PAPAIA:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
- <rename>
  <version>PAPAIAfile1.0</version>
  <prefix>ULFL55172-10</prefix>
  <date>2002-10</date>
  <image_format>tif</image_format>
  <pages_per_image>1</pages_per_image>
  <sorting_mode>by index</sorting_mode>
  <dir>C:\GUARDAR IMAGENS\Artis\ULFL55172-10_p</dir>
  <image_numbering_sel>0</image_numbering_sel>
  <page_numbering_sel>0</page_numbering_sel>
  <formula>%c_%d_%i_%p_t0</formula>
  <auto_end />
- <page_block>
  <type>Custom</type>
  <typeID>1</typeID>
  <start>1</start>
  <end>2</end>
  <date>2002-10</date>
  <custom>capa</custom>
  <file>C:\GUARDAR IMAGENS\Artis\ULFL55172-10_p\ULFL55172-10_2002-10_0001_capa_t0.tif</file>
  <file>C:\GUARDAR IMAGENS\Artis\ULFL55172-10_p\ULFL55172-10_2002-10_0002_capa_t0.tif</file>
</page_block>
- <page_block>
  <type>Custom</type>
  <typeID>1</typeID>
  <start>1</start>
  <end>2</end>
  <date>2002-10</date>
  <custom>b</custom>
  <file>C:\GUARDAR IMAGENS\Artis\ULFL55172-10_p\ULFL55172-10_2002-10_0003_b_t0.tif</file>
  <file>C:\GUARDAR IMAGENS\Artis\ULFL55172-10_p\ULFL55172-10_2002-10_0004_b_t0.tif</file>
</page_block>
</name_block>
```

Figura 16

8) Resultado final da nomeação e estruturação de ficheiros efectuado pelo programa PAPAIA:

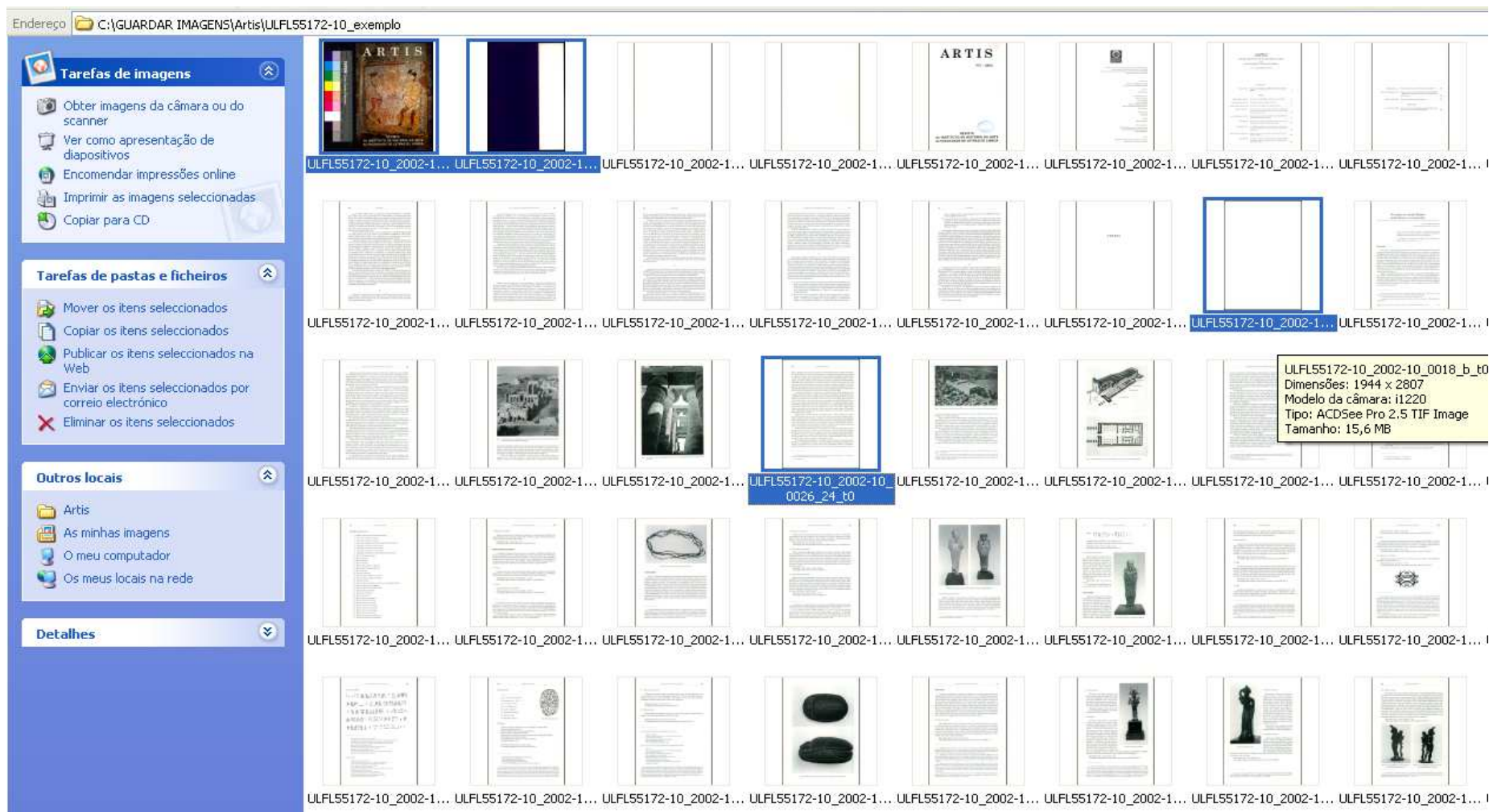


Figura 17

ANEXO 6

CONTROLO DE QUALIDADE DAS IMAGENS

No sentido de haver um efectivo controlo de qualidade das imagens digitalizadas, foi definida uma tabela com os nomes escalonados das pessoas envolvidas no projecto: estagiária, voluntárias e funcionária da biblioteca, segundo a premissa de que quem digitalizava uma revista não deveria fazer o controlo de qualidade da mesma.

Controle de qualidade						
	Código de barras	Sónia	Ana Paula	Marina	Raquel	Graça
A r t i s	ULFL55172-10					x
	ULFL55172-20					x
	ULFL55172-30			x		
	ULFL55172-40			x		
	ULFL55172-50				x	
	ULFL55172-60				x	
P h i l o s o p h i c a	ULFL14278-10	x				
	ULFL14278-20	x				
	ULFL14278-30		x			
	ULFL14278-40		x			
	ULFL14278-50					x
	ULFL14278-60					x
	ULFL14278-70			x		
	ULFL14278-80			x		
	ULFL14278-100				x	
	ULFL14278-110				x	
	ULFL14278-120	x				
	ULFL14278-150	x				
	ULFL14278-160		x			
	ULFL14278-180		x			
R e v i s t a d a F a c u l d a d e	ULFL14278-190					x
	ULFL14278-210					x
	ULFL14278-220			x		
	ULFL13401-40			x		
	ULFL13401-50				x	
	ULFL13401-60				x	
	ULFL13401-670	x				
	ULFL13401-250	x				
	ULFL13401-260		x			
	ULFL13401-280		x			
	ULFL13401-290					x
	ULFL13401-300					x
	ULFL13401-310			x		
	ULFL13401-320			x		
	ULFL13401-330				x	
	ULFL13401-340				x	
	ULFL13401-350	x				
	ULFL13401-360	x				
	ULFL13401-370		x			
	ULFL13401-380		x			
	ULFL13401-390					x
	ULFL13401-410					x
	ULFL13401-420			x		
	ULFL13401-430			x		
ULFL13401-440				x		
ULFL13401-450				x		
ULFL13401-460	x					
ULFL13401-470	x					
ULFL13401-480		x				
ULFL13401-490		x				
ULFL13401-500					x	
ULFL13401-510					x	
ULFL13401-520			x			
ULFL13401-530			x			

cfaria

31-03-2009 18:58

cfaria

31-03-2009 18:58

Figura 1 – Escala do controlo de qualidade

[illegible]

De facto, através deste método detectaram-se alguns erros que se não tivessem sido encontrados colocavam em causa o nosso projecto. O exemplo seguinte é representativo da eficácia do método que escolhemos. A verificação teria de ser feita por três pessoas até ser considerada como definitiva, no caso do

Controle de qualidade (TIFF)		FOLHA DE OBRA N.º 2	
IDENTIFICAÇÃO			
LOTE:	ULFL 55172-20	N.º IMAGENS:	332
QUALIDADE DA IMAGEM DIGITAL			
1. Nome da pasta	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
2. Margens certas	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
3. Imagens seguem sequência da obra	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
4. Orientação correcta	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
5. Todas p. digitalizadas	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
6. Cunha de cor (1.ª e última imagem)	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
7. Nitidez da imagem	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
PAPALIA			
8. Nomeação de cada imagem	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
9. N.º página coincide com n.º obra	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
METADADOS			
10. Preenchidos na totalidade	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
11. Correctamente preenchidos	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
GRAVAÇÃO			
12. Disco externo	S <input checked="" type="checkbox"/>	N	<input type="checkbox"/>
ANOTAÇÃO DE FALHAS DETECTADAS			
<p>Pág 273 fora de ordem</p>			
VERIFICAÇÃO			
1º. Por Luis Fernandes	Em 11/5/2009	Aceite <input checked="" type="checkbox"/>	Recusado <input type="checkbox"/>
2º. Por Graça Lopes	Em 19/6/2009	Aceite <input type="checkbox"/>	Recusado <input checked="" type="checkbox"/>
3º. Por	Em	Aceite <input type="checkbox"/>	Recusado <input type="checkbox"/>

Quando se regista uma falha em algum número, esta tem de ser corrigida e esse número é enviado de novo para o controlo de qualidade até estar completamente correcto e ser dado como definitivo.

ANEXO 7

O PROCESSO DE DISPONIBILIZAÇÃO – ContentE

1) Abrir uma obra analítica no programa ContentE

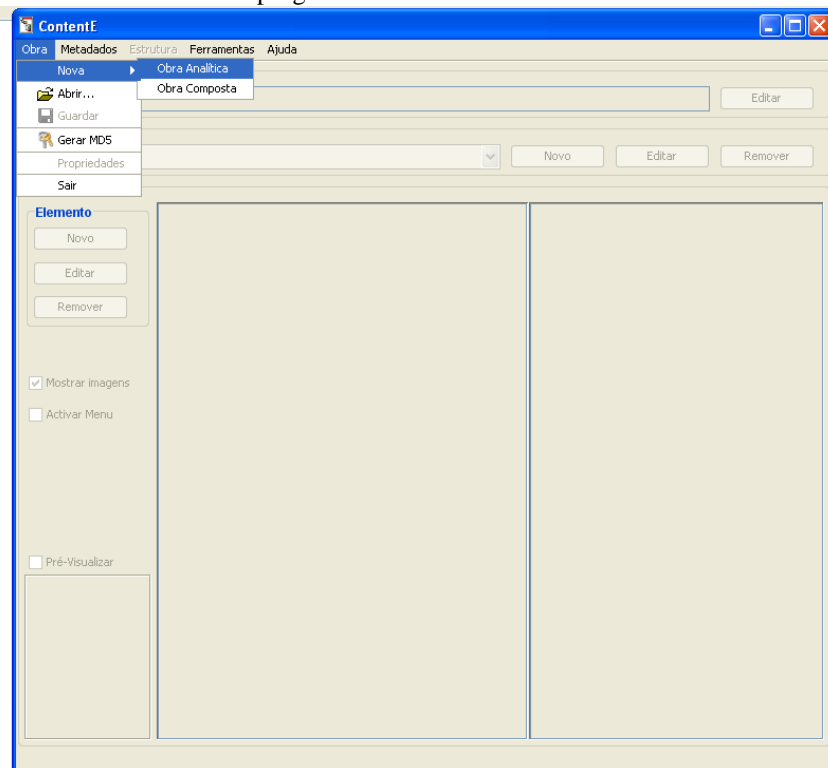


Figura 1

2) Importar a pasta com as imagens TIFF

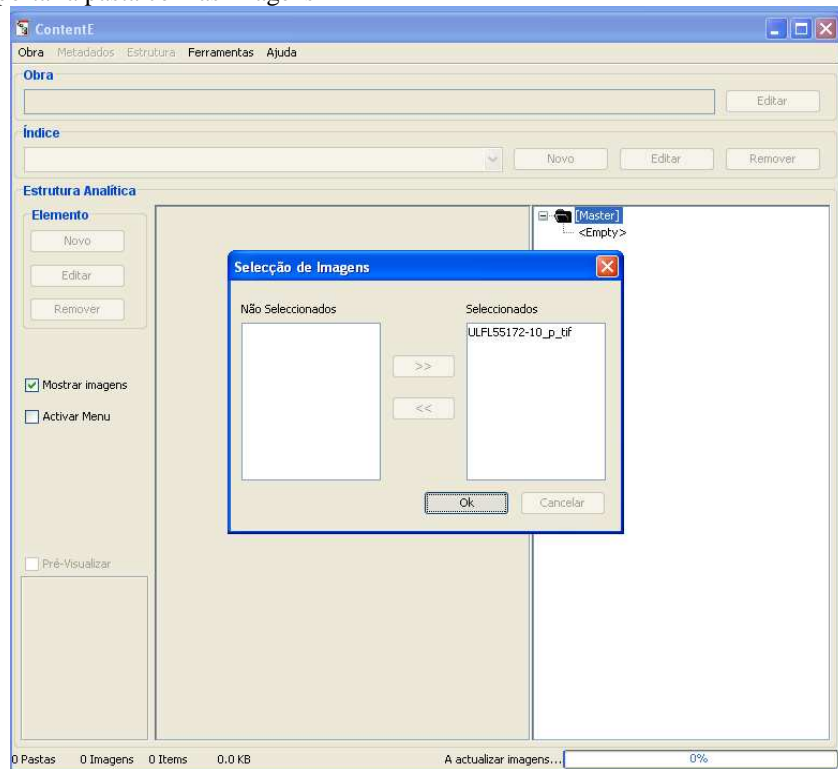


Figura 2

Como podemos verificar no exemplo seguinte, as imagens que importamos são depositadas no separador do lado direito.

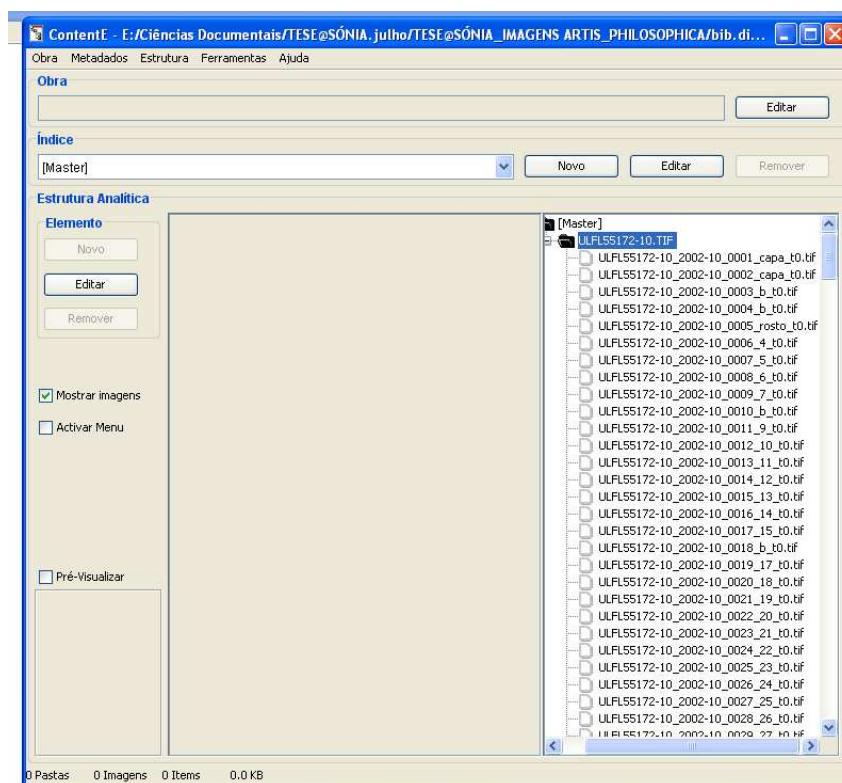


Figura 3

Esta versão do ContentE permite a conversão de ficheiros TIFF em formato JPEG e PDF. Uma vez que as imagens que iremos disponibilizar terão de ser em formato JPEG, esta conversão tem de ser efectuada.

a) Conversão de TIFF em JPEG

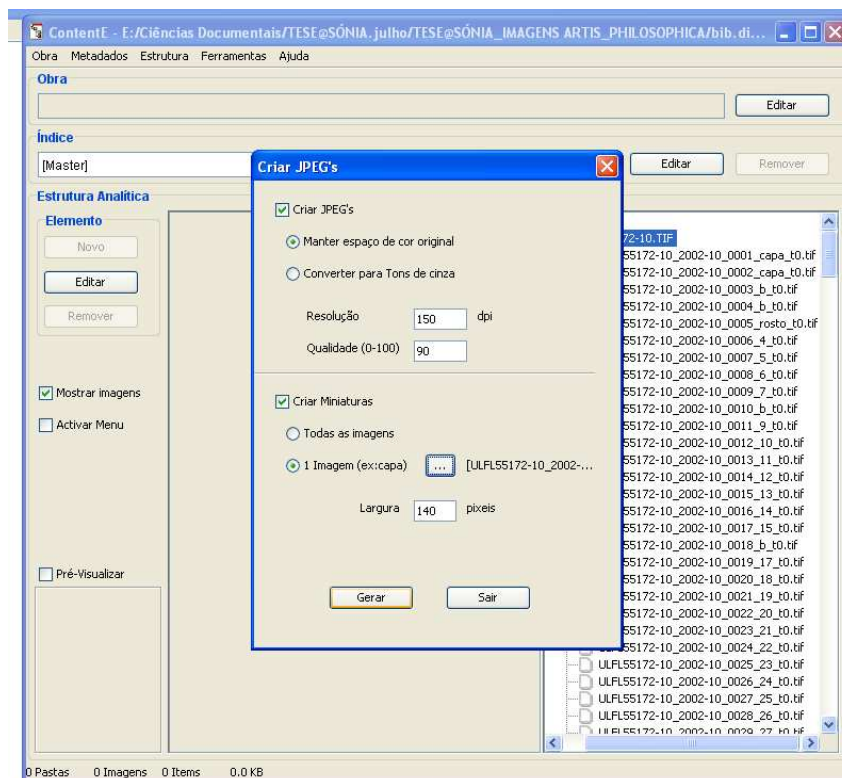


Figura 4

b) Conversão de TIFF em PDF:

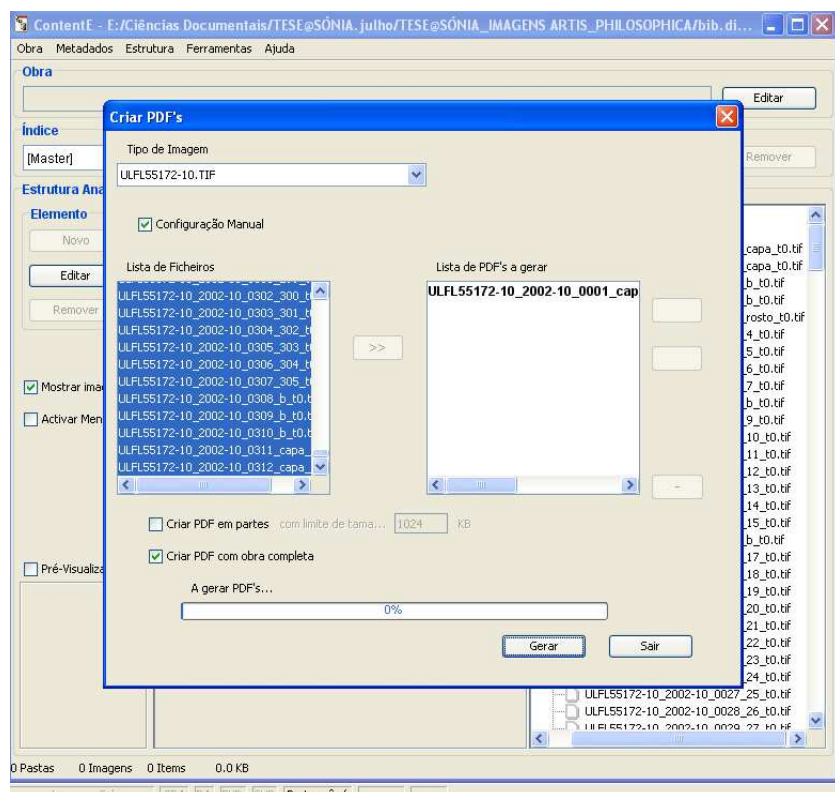


Figura 5

Podemos anexar à revista que estamos a estruturar o seu registo bibliográfico, basta importá-lo para o programa, como podemos visualizar no exemplo seguinte:

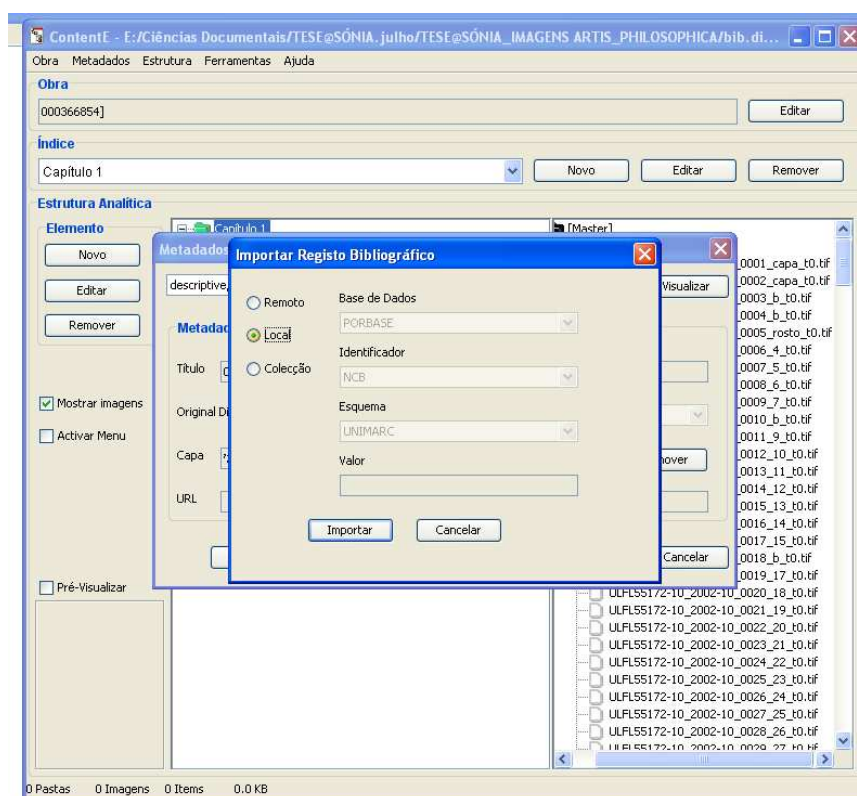


Figura 6

Continuação da colocação do registo bibliográfico:

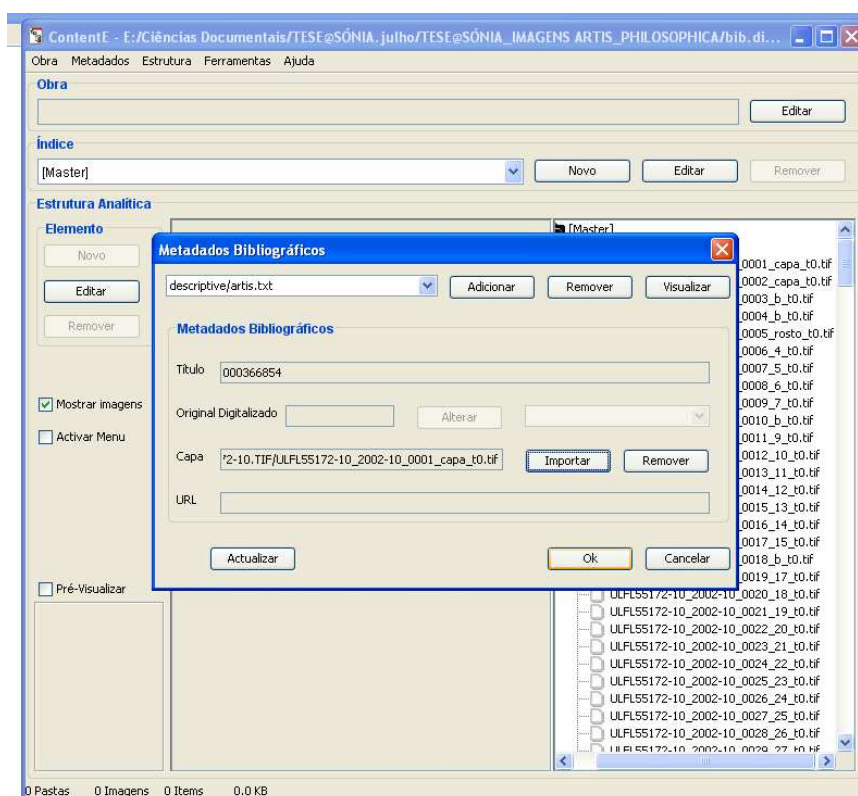


Figura 7

O próximo passo é criar o índice da respectiva obra:

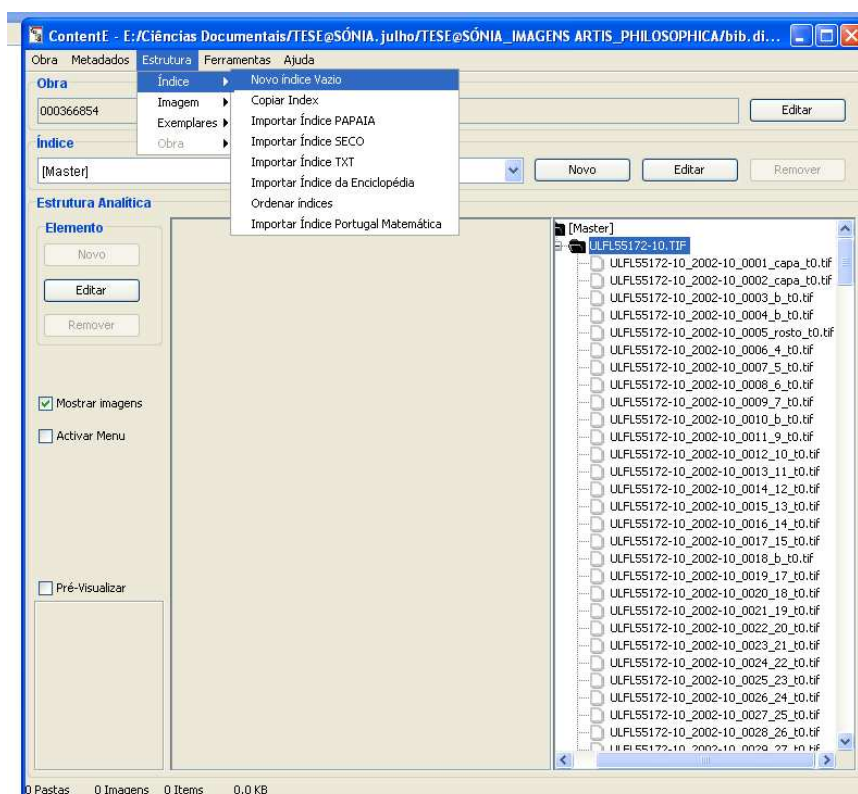


Figura 8

Escolhemos o nome que queremos dar ao tipo de nó assim como o tipo de numeração:

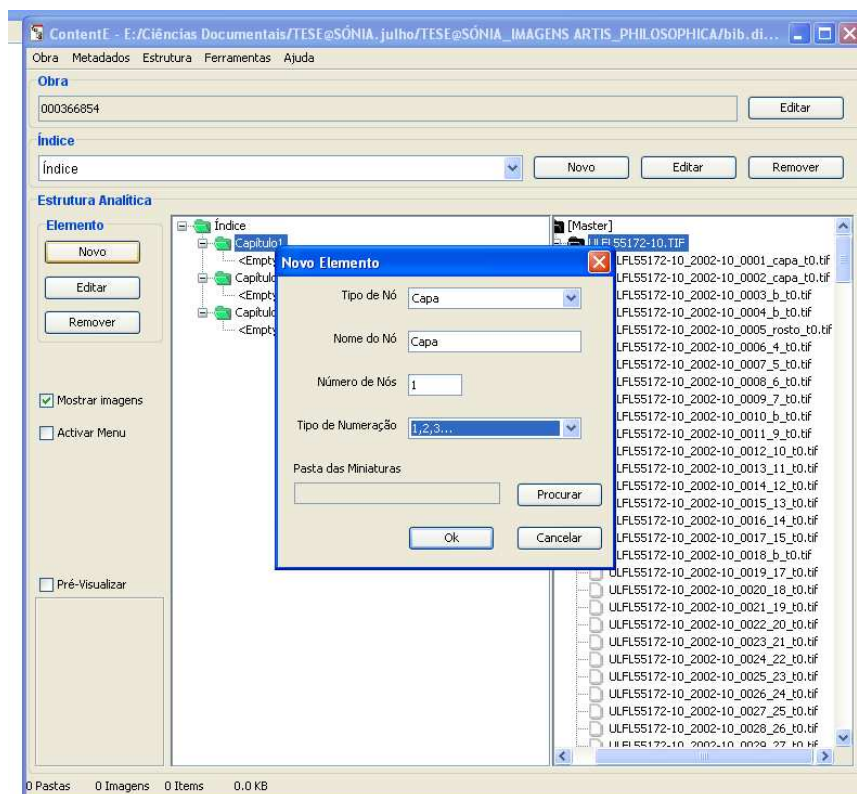


Figura 9

Exemplo da evolução da construção de um índice:

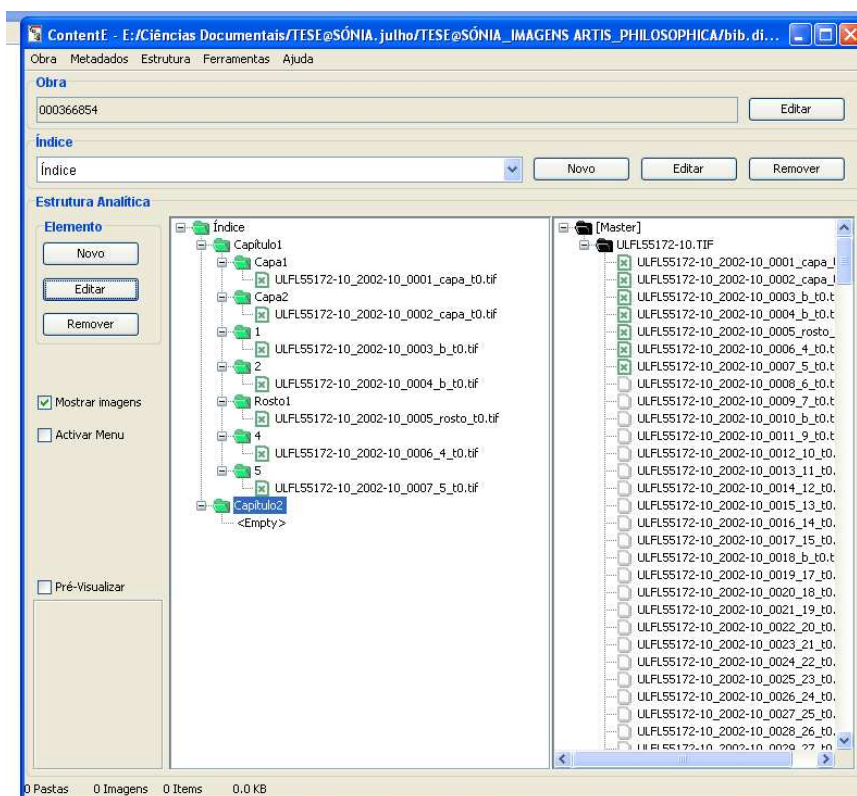


Figura 10

Após a criação do índice podemos agregar ao nosso trabalho uma “folha de estilo” que personalizará a página *Web* onde os conteúdos serão disponibilizados:

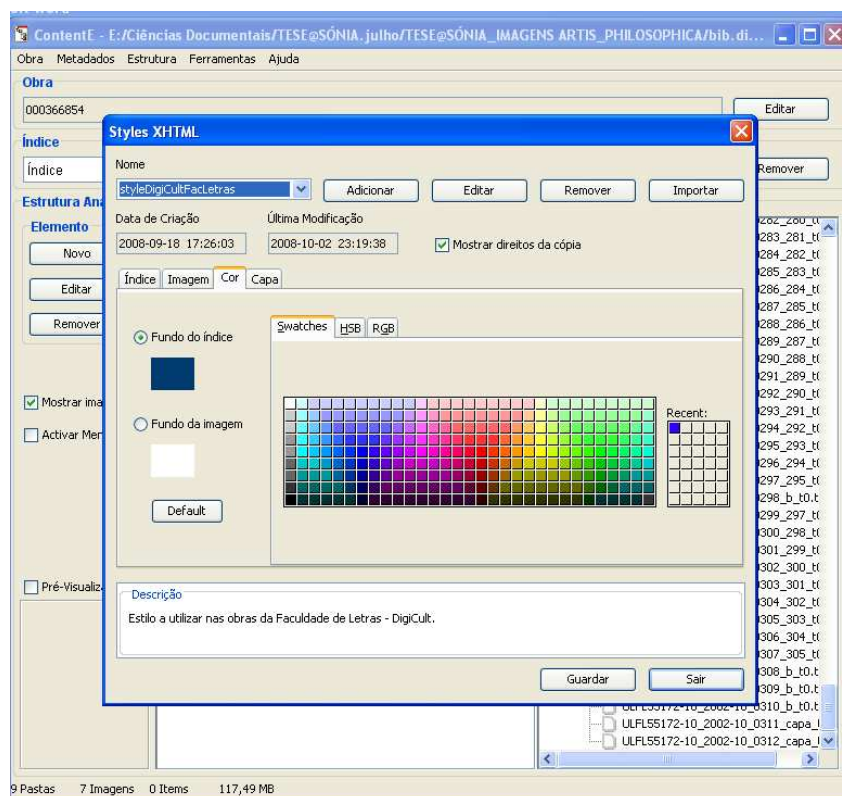


Figura 11

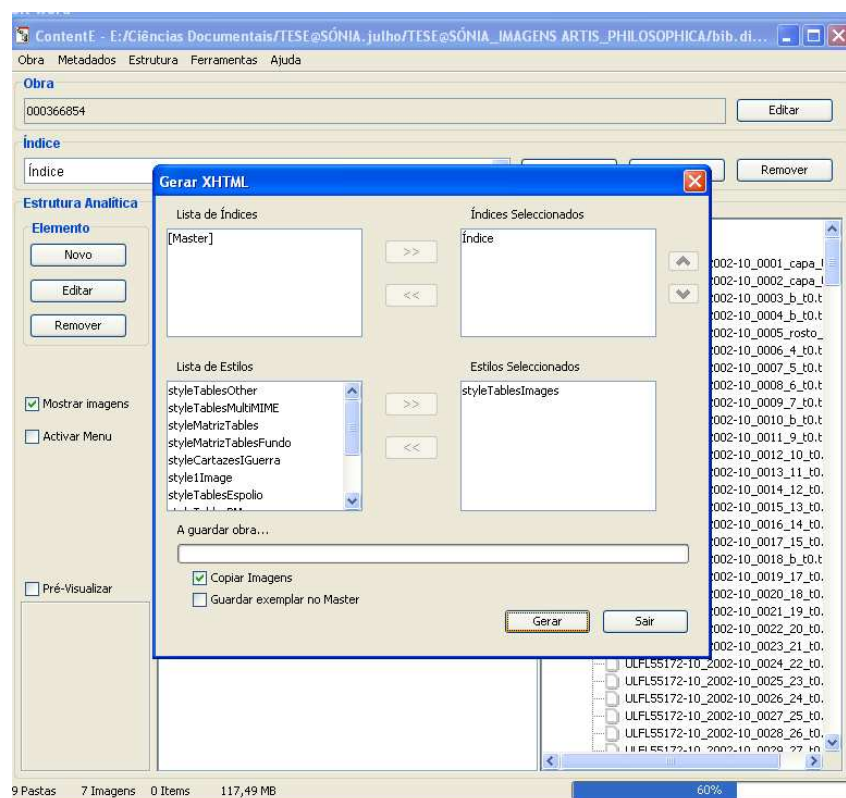


Figura 12